Ⓔ

PRENOLIN: International Benchmark on 1D Nonlinear Site-Response Analysis—Validation Phase Exercise

by Julie Régnier, Luis-Fabian Bonilla, Pierre-Yves Bard, Etienne Bertrand, Fabrice Hollender, Hiroshi Kawase, Deborah Sicilia, Pedro Arduino, Angelo Amorosi, Dominiki Asimaki, Daniela Boldini, Long Chen, Anna Chiaradonna, Florent DeMartin, Ahmed Elgamal, Gaetano Falcone, Evelyne Foerster, Sebastiano Foti, Evangelia Garini, George Gazetas, Céline Gélis, Alborz Ghofrani, Amalia Giannakou, James Gingery, Nathalie Glinsky, Joseph Harmon, Youssef Hashash, Susumu Iai, Steve Kramer, Stavroula Kontoe, Jozef Kristek, Giuseppe Lanzo, Annamaria di Lernia, Fernando Lopez-Caballero, Marianne Marot, Graeme McAllister, E. Diego Mercerat, Peter Moczo, Silvana Montoya-Noguera,[*] Michael Musgrove, Alex Nieto-Ferro, Alessandro Pagliaroli, Federico Passeri, Aneta Richterova, Suwal Sajana, Maria Paola Santisi d'Avila, Jian Shi, Francesco Silvestri, Mahdi Taiebat, Giuseppe Tropeano, Didrik Vandeputte, and Luca Verrucci

Abstract This article presents the main results of the validation phase of the PRENOLIN project. PRENOLIN is an international benchmark on 1D nonlinear (NL) site-response analysis. This project involved 19 teams with 23 different codes tested. It was divided into two phases; with the first phase verifying the numerical solution of these codes on idealized soil profiles using simple signals and real seismic records. The second phase described in this article referred to code validation for the analysis of real instrumented sites.

This validation phase was performed on two sites (KSRH10 and Sendai) of the Japanese strong-motion networks KiK-net and Port and Airport Research Institute (PARI), respectively, with a pair of accelerometers at surface and depth. Extensive additional site characterizations were performed at both sites involving *in situ* and laboratory measurements of the soil properties. At each site, sets of input motions were selected to represent different peak ground acceleration (PGA) and frequency content. It was found that the code-to-code variability given by the standard deviation of the computed surface-response spectra is around 0.1 (in log10 scale) regardless of the site and input motions. This indicates a quite large influence of the numerical methods on site-effect assessment and more generally on seismic hazard. Besides, it was observed that site-specific measurements are of primary importance for defining the input data in site-response analysis. The NL parameters obtained from the laboratory measurements should be compared with curves coming from the literature. Finally, the lessons learned from this exercise are synthesized, resulting also in a few recommendations for future benchmarking studies, and the use of 1D NL, total stress site-response analysis.

Electronic Supplement: Table of the participants and codes used, and figures of localization of soil sampling and comparisons of response spectra and spectral ratios for data and synthetics.

Introduction

In seismology and earthquake engineering, site effects are widely recognized as an important factor for (mainly) amplifying the resulting surface ground motion. Those site effects are spatially variable depending on the local geomorphology and mechanical properties of the soil; they may vary from one event to the other as the site response to seismic loading is nonlinear (NL) during strong ground motion

---

*Now at Department of Civil Engineering, Universidad EAFIT, Medellín, Colombia.

(e.g., Seed and Idriss, 1969; Vucetic and Dobry, 1991; Ishibashi and Zhang, 1993; Yu *et al.*, 1993; Elgamal *et al.*, 1995; Zeghal *et al.*, 1995; Gunturi *et al.*, 1998; Bonilla *et al.*, 2005; Amorosi *et al.*, 2016).

Site-specific analysis of the site response involving its NL soil behavior is still very challenging. In low-seismicity areas, the lack of strong ground motion recordings limits empirical evaluations. To overcome this limitation, numerical simulations are of primary interest because they allow for simulating strong ground motions beyond the available recordings. In engineering practice, those analyses would, in most of the cases, be limited to methods involving the use of linear or equivalent linear (EQL) methods in a 1D site configuration.

As mentioned in the preceding companion paper (Régnier, Bonilla, *et al.*, 2016), previous comparative tests were performed to study 1D NL site response. The first blind tests performed in the late 1980s/early 1990s, on the Ashigara Valley (Japan) and Turkey Flat (California) sites were very instructive in the linear domain, because of the lack of strong-motion recordings at that time. The 2004 Parkfield earthquake, which produced $0.3g$ at Turkey Flat site, was the opportunity to launch a new benchmarking exercise for 1D NL codes. This also considered a few other sites with vertical array data and large enough ground motion (La Cienega, California; the KGWH02 KiK-net site, Japan; and a site in Lotung, Taiwan). The results reported in Kwok *et al.* (2008), Stewart (2008), and Stewart and Kwok (2008, 2009) emphasized the importance of the needed *in situ* measurements and of the actual way these codes are used. The origin of the significant mismatch between records and predictions has been attributed to incorrect velocity profiles (despite the redundant borehole measurements), to deviations from 1D geometry (nonhorizontal soil layers), and deficiencies in the constitutive models (unsatisfactory match to the actual degradation curves). The 1D codes used for these tests remain, however, extensively used in routine engineering practice for site-response estimates, and various developments have been carried out to implement new, or updated, constitutive models. It is therefore needed to repeat such benchmarking activities, notably in other parts of the world, which may have different engineering practices, and were not involved in the previous comparison exercises (which are also always good for young scientists and engineers who never had such a benchmarking experience).

The objective of this two-phase (verification and validation) PRENOLIN exercise is to understand the variability associated with the implementation of the NL soil behavior in numerical simulations, and to assess the resulting uncertainties. It was decided to devote the calculations on simple cases focusing the numerical implementation of the NL soil behavior (rheology and soil parameters) to be as close as possible to the standard engineering practice.

In this work, we evaluate 1D wave propagation of *SH* waves (only one component of motion) with vertical incidence and assuming no pore pressure effects (total stress analysis). These three assumptions mentioned above are not a sensitive issue when dealing with the verification case (although realistic cases were selected to be close to the true physical processes). However, when dealing with validation and comparison with real data, they may have very strong impacts and for most sites and input motions they might be violated to various degrees. We choose our sites for the validation to minimize the impact of those assumptions.

The verification phase helped create a synergy between the participants and the organizing teams. We defined a common vocabulary for the implementation of the calculations (as the NL communities from different areas of the world may have different practices and a different understanding of the same words). By analyzing the whole set of NL simulations, we found that the code-to-code variability increases with the shear-strain level. However, even in the worst case corresponding to large loading and strain levels exceeding 1%, it remained lower than the single-station random variability of ground-motion prediction equation (GMPE) σ-values for peak ground acceleration (PGA). Given the scatter in the NL results, we thus concluded that a realistic analysis should use more than one constitutive model to incorporate at least partially the epistemic uncertainty in site-response computations. It was also found that, to reduce the epistemic uncertainty, which is partially accounted for by the code-to-code variability, one may need to precisely describe specific input parameters, especially the soil shear strength profile. In addition, the variability between codes is considerably reduced when they all used the same loading and unloading process (Masing rules or non-Masing rules, referred to as damping control models) (Régnier, Bonilla, *et al.*, 2016).

To keep this exercise relatively simple, it was decided from the beginning not to deal with the problem of pore pressure effects (e.g., Zeghal *et al.*, 1995; Elgamal *et al.*, 1996). We acknowledge that this hypothesis might not be realistic in saturated soils subjected to strong-motion cyclic loads. One of the chosen Port and Airport Research Institute (PARI) sites (Onahama) may have experienced cyclic mobility during the 11 March 2011 M_w 9 Tohoku-Oki earthquake (Roten *et al.*, 2013). However, dealing with NL site response increases the number of soil parameters to consider, not only with respect to traditional linear estimates but also in relation to the complexity of the constitutive model used to describe the soil NL behavior. Because soil NL behavior, even in the absence of pore pressure effects, remains a challenge, the main effort of the validation exercise was focused on total stress analysis only, disregarding the simulation of excess pore pressure generation. To verify effects of water pressure buildup on the recordings acceleration time history, some teams used site-response analyses with excess pore-water pressure generation to compare with the total stress cases and are shown in Ⓔ Figures S6 and S7 (available in the electronic supplement to this article). The recommendations of this report are only for total stress site-response analyses.

A detailed presentation of the organization of the project and the participant teams was done in the previous paper related to the verification phase (Régnier, Bonilla, *et al.*, 2016). The table of the participants with description of the methods is thus provided here only in Ⓔ Table S1.

The objectives of this article are to share the experiences on the validation exercise concerning (1) the analysis of data from laboratory and *in situ* tests data to define the input soil parameters for the simulations, (2) the processing and selection of the seismic input motions, (3) the calculations of the NL site response, and finally (4) to quantify the epistemic uncertainty for 1D NL site-response analysis on real sites, both in terms of code-to-code variability and code-to-data distance. All the results and cases of the study used in this benchmark are available in a dedicated website (see Data and Resources).

Target Sites

Selection Criteria

The selection of sites was performed on strong-motion databases involving vertical arrays so that the empirical soil column response (often called transfer function) can be calculated. Considering the hypothesis of the numerical methods and the objective to implement NL soil behavior, the sites were selected on the basis of the following requirements:

1. availability of both strong- and weak-motion recordings;
2. plausibility of 1D geometrical soil configuration, that is, satisfactory agreement between numerical and empirical site responses in the linear/weak-motion range;
3. the depth of downhole sensor must be less than 300 m; and
4. the possibility to perform complementary investigations in the immediate proximity of the site.

To fulfill the first and second criteria, sites that recorded at least two earthquakes with PGAs higher than 50 cm/s^2 at the downhole sensor were selected. Only the KiK-net site configurations identified as fulfilling the 1D criteria proposed by Thompson *et al.* (2012) and Régnier *et al.* (2014) were considered. In addition, a visual comparison between the numerical and empirical site-response curves was performed and a special attention was given to the matching of fundamental resonance frequency.

The fourth criterion also constrained the site selection, as the nearby urbanization may prevent the drilling of new boreholes or the ability to perform new surface measurements.

Dataset

Presentation. The site selection was done on the KiK-net and PARI networks. KiK-net is composed of 688 stations, with high-quality surface and downhole digital three-component accelerometers. Among the KiK-net sites, 668 are characterized with shear- and compressive-wave-velocity profiles. These velocity profiles were obtained from downhole or *PS*-logging measurements (depending on the site).

Most of the borehole sensors are located between 100 and 200 m depth. Two-thirds of the sites have a $V_{S30} < 550$ m/s. In the National Earthquake Hazards Reduction Program and Eurocode 8 regulation, sites with $V_{S30} < 800$ m/s are classified as sites prone to site effects, which confirm that the KiK-net database is very interesting for the analysis of these phenomena. The PARI sites are much shallower than the KiK-net sites: the downhole sensor depth is around 10 m and the corresponding V_S profile is also available.

Data Processing. More than 46,000 (six-component) recordings were analyzed beforehand (Régnier *et al.*, 2013) to derive the empirical site response (i.e., the transfer function from downhole to surface) at the 688 sites. In addition, on 668 sites with available V_S profiles, the numerical linear site response was also calculated on the basis of the velocity profile provided in the KiK-net database (see Data and Resources). On the PARI network, only two sites were analyzed: Sendai (30 earthquake recordings) and Onahama (42 earthquake recordings).

The empirical site response is usually evaluated using a spectral ratio between simultaneous recordings on sediments and on a nearby rock site (the so-called reference site). When this technique is applied, the main issue to overcome is the selection of a reliable reference site. The reference site should not amplify seismic waves and should be close enough to the studied site so that the travel path from the seismic source remains equivalent for both sites.

Vertical arrays of accelerometers overcome the reference-site issue. Indeed, the downhole station located at depth represents the reference station. Thus, for each KiK-net (and PARI) sites and each earthquake recording, the borehole Fourier spectral ratio (BFSR) were calculated. BFSR is the ratio between the Fourier spectra of the horizontal components recorded at the surface and the corresponding ones at depth. Yet we acknowledge that the use of downhole records introduces an additional difficulty in numerical modeling due to the contamination of the control motion by the downgoing wavefield, which is sensitive both to the details of velocity and damping soil profile, and to the complexity of the incoming wavefield (various types of body waves with multiple incidence angles, together with possibility of surface waves; see Bonilla *et al.*, 2002; Régnier *et al.*, 2014).

Before evaluating the BFSR, a specific data processing procedure was applied that consisted of removing the mean, applying a tapering Hanning window on 2% of the signal, noncausal filtering between 0.1 and 40 Hz, fast Fourier transform calculation and a Konno–Omachi smoothing (with $b = 40$) before performing the surface to downhole spectral ratio. The linear site response was obtained by calculating the geometric average of all recordings with a PGA at the surface below 25 cm/s^2.

The empirical site-response curves were compared with the equivalent numerical ones (BFSR$_{num}$). The numerical site responses (BFSR$_{num}$ and outcrop Fourier spectral ratio OFSR$_{num}$, indicating, respectively, the transfer functions that
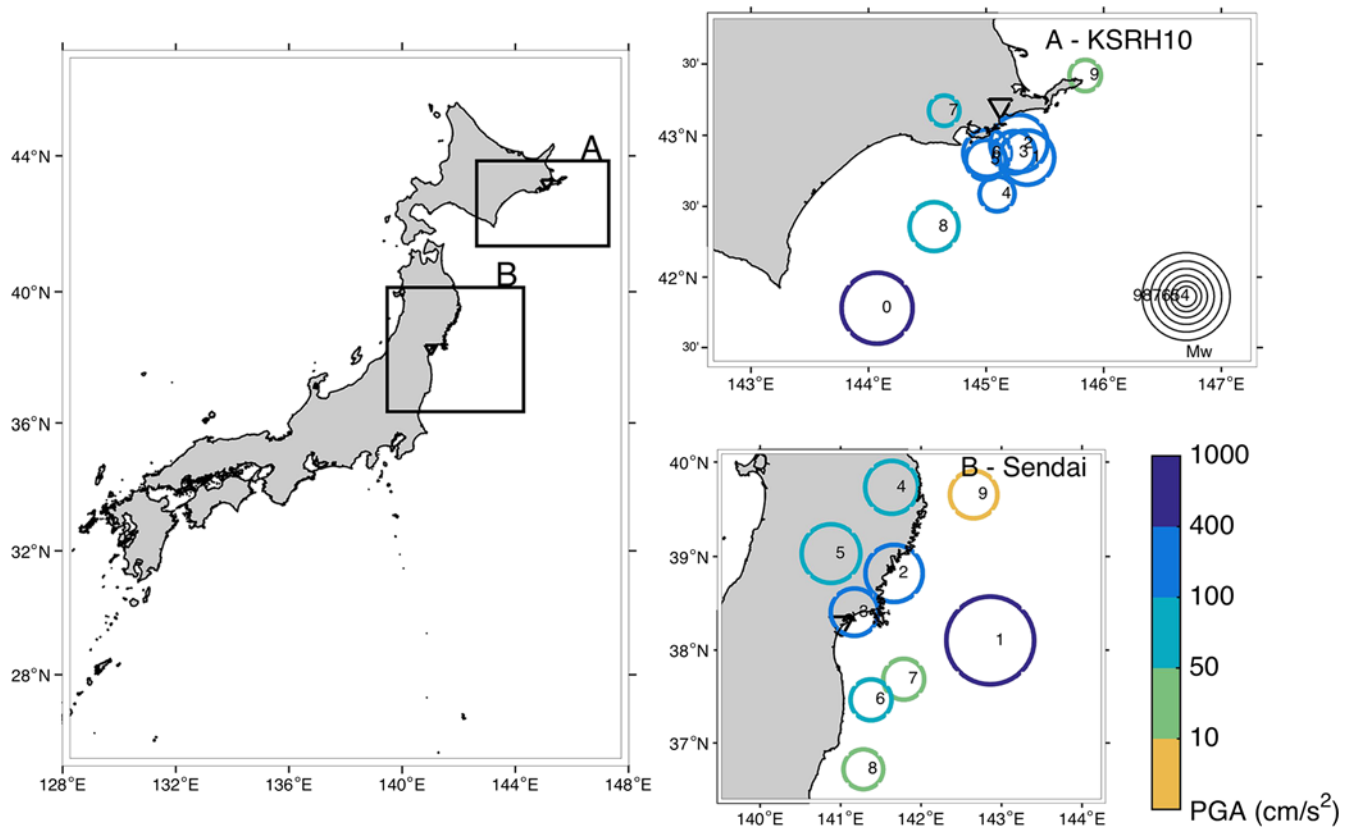
Figure 1. Location of the two selected sites for the validation phase with location of the epicenter of the selected events at each site according to their magnitude and surface peak ground acceleration (PGA; cm/s² at the station). The color version of this figure is available only in the electronic edition.

have the reference at the downhole and at the rock outcrop, respectively) were computed using a Haskell–Thomson 1D linear method (Thomson, 1950; Haskell, 1953). For the calculation of the transfer function, we also added a Konno–Ohmachi smoothing ($b = 40$) that was applied directly to the transfer function curve (because it is the direct result from the Haskell–Thomson method) and the same frequency sampling as the one dealing with the analysis of earthquake recordings was used.

The soil parameters that are required to compute the numerical transfer functions are the shear-wave velocity profile V_S, the density profile, and the quality factor profile. In the KiK-net (and PARI) database, only the V_S profiles are available. For the density, a constant value along the profile equal to 2000 kg/m³ was used. The quality factor (Q) was directly derived from the V_S value following the rule of thumb scaling: $Q = V_S/10$ (Olsen *et al.*, 2003) used by many authors when no measurement is available from laboratory data measurements (see the From *In Situ* and Lab Data to Input Parameters section). Other models for the low-strain attenuation also could have been used, as proposed in Darendeli (2001) and Menq (2003).

Selected Sites

Five KiK-net sites (FKSH14, IBRH13, IWTH04, KSRH10, and NIGH13) passed the selection criteria, together with the two PARI sites that were initially chosen. Four KiK-net sites were removed for various reasons: liquefaction susceptibility (FKSH14), rocky geology (IBRH13), mountainous environment (IWTH04), and insufficient nonlinearity (NIGH13). A detailed study on the effects of the topography and nonhorizontal layering on waveforms and transfer functions of KiK-net sites can be found in De Martin *et al.* (2013).

The three remaining sites, that is, KSRH10, Onahama, and Sendai sites were therefore selected for further *in situ* investigations for the purpose of the validation phase. Later on, for the Onahama site, it was found that the soil was susceptible to liquefaction (Roten *et al.*, 2013) and clearly show 2D/3D site configuration. Consequently, the calculations performed on this site will not be presented in the present article, although they were part of the validation phase.

The locations of KSRH10 (Hokkaido region) and Sendai (Tohoku region) are illustrated in Figure 1.

According to the initial available geotechnical data, KSRH10 is mainly composed of clayey soil, whereas Sendai site is composed of sandy soil. KSRH10 site is a deep sedimentary site with 40 m of low-velocity soil layers; the downhole sensor is located at a depth of 255 m (Fig. 2a). The site is located on the lower plateaus with about 30 m in elevation along the right bank (southern) side of the upper Anebetsu River. The soil column consists of recent Younger Volcanic
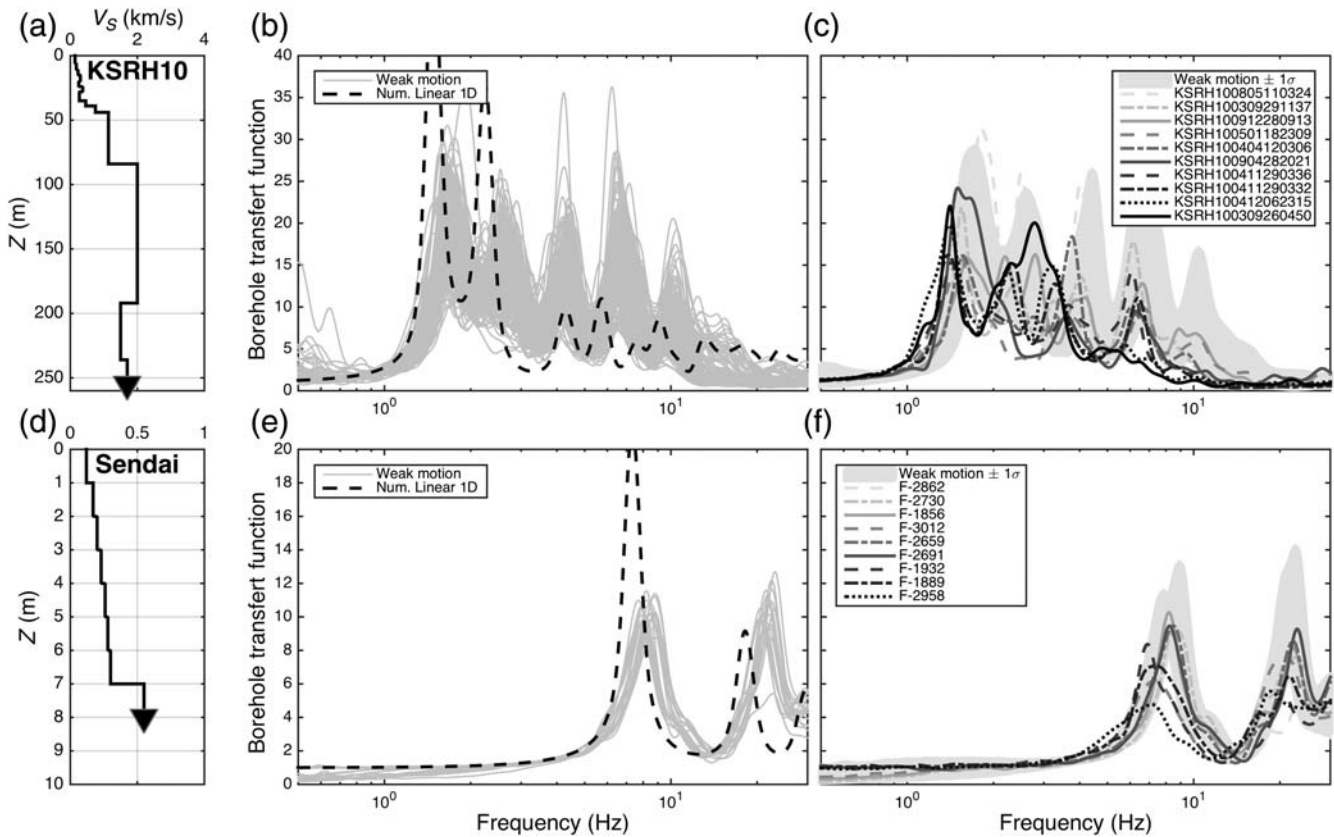
Figure 2.   (a,d) The initial V_S profile available at the selected sites for KSRH10 and Sendai. (b,e) The comparison between 1D linear site response computed with empirical site response (surface to within motion spectral ratio) calculated with weak motion (PGA at the surface lower than 25 cm/s²) for KSRH10 and Sendai sites. (c,f) The comparison of the empirical transfer function of weak motions with the strongest ones recorded at KSRH10 and Sendai, respectively.

Ash deposits until 5 m at depth, followed by volcanic and Tuffaceous sand until 40 m and underlie by an alternation of sandstone and shale.

The Sendai site is a shallow site with 7 m of soft soil deposits and with the downhole sensor located at 10.4 m at depth. According to the Shogama 1:50,000 geological map, the site is in a large flat valley covered by beach ridge deposits (Holocene) consisting of gravel and sand. These surface deposits are underlain by the Pliocene, Geba Formation, forming the northern and eastern hills and consist of gravel stone, sandstone, tuff, tuffaceous siltstone, and lignite.

As illustrated in Figure 2b,e, empirical weak-motion BFSR (surface PGA lower than 25 cm/s²) and the linear numerical 1D site response (dashed line) exhibit a satisfactory similarity, especially for Sendai site. The shallow Sendai site is characterized by a high resonance frequency around 8.2 Hz, whereas the thicker KSRH10 site is characterized by a lower fundamental resonance frequency of 1.7 Hz; slight differences can be seen however between observations and simulations as to the frequencies of the first 2 peaks. The numerical simulations (dashed lines) provide site-response amplitude much higher at the first 2 frequency peaks and above 15 Hz, whereas it is lower for the third and fourth peaks. This first comparison shows that the sites are close

to a 1D site configuration, but exhibit a more complex behavior than those predicted for a simple soil column subjected to pure vertically incident plane S waves.

When the sites were selected for the validation, one of the requirements was that the sites exhibit some NL soil behavior for one or several recordings. In Figure 2c,f, the BFSR for weak motion is compared to the BFSR calculated from motions with large PGA at the surface. For KSRH10, NL soil behavior is significant for the input motions with PGA greater than 47 cm/s² at the downhole station (outside the average ± standard deviation area) and it is even greater for the strongest events (KSRH100411290332 with PGA equal to 81 cm/s² and KSRH100309260450 with PGA equal to 110 cm/s²; Table 1). For Sendai, NL soil behavior is significant when the downhole PGA exceeds 46 cm/s² (outside the average ± standard deviation area) and it is even greater for the strongest event (F-2958 with PGA equal to 252 cm/s²).

Selection of the Input Motions

A selection of 10 and 9 input motions for KSRH10 and Sendai, respectively, was performed among the available earthquake recordings. Their epicenters, magnitudes, and

Table 1
Selected Event Characteristics (Peak Ground Acceleration [PGA] for East–West [EW] Components)

	Earthquake	Code Name	F_s (Hz)	M_w	Depth (km)	D_{epi} (km)	$PGA_{downhole}$ (cm/s²)	$PGA_{surface}$ (cm/s²)	F_c (Hz)
KSRH10	**KSRH100309260450**	**TS-0-K**	**200**	**8**	**42**	**180**	**110**	**558**	**6.6**
	KSRH100411290332	**TS-1-K**	**200**	**7.1**	**48**	**32**	**81**	**319**	**4.8**
	KSRH100412062315	**TS-2-K**	**200**	**6.9**	**46**	**44**	**69**	**386**	**4.2**
	KSRH100411290336	TS-3-K	200	6	46	37	64	199	6.0
	KSRH100404120306	**TS-4-K**	**200**	**5.8**	**47**	**43**	**27**	**162**	**5.3**
	KSRH100904282021	TS-5-K	100	5.4	38	69	25	163	4.0
	KSRH100501182309	TS-6-K	200	6.4	50	38	25	125	6.7
	KSRH100912280913	TS-7-K	100	5	85	39	9	58	6.5
	KSRH100805110324	TS-8-K	100	5.1	88	63	8	46	6.2
	KSRH100309291137	**TS-9-K**	**200**	**6.5**	**43**	**105**	**7**	**54**	**4.6**
Sendai	**F-2958**	**TS-1-S**	**100**	**9**	**23.7**	**163**	**252**	**481**	**6.9**
	F-1889	**TS-2-S**	**100**	**7.1**	**72**	**81**	**62**	**244**	**9.0**
	F-1932	TS-3-S	100	6.4	11.9	19	61	208	10.3
	F-2691	TS-4-S	100	6.8	108.1	169	25	89	7.1
	F-3012	**TS-5-S**	**100**	**5.9**	**30.7**	**96**	**25**	**72**	**7.5**
	F-2659	TS-6-S	100	7.2	7.8	83	35	82	7.7
	F-1856	TS-7-S	100	5.9	41.2	95	12	32	7.8
	F-2862	**TS-8-S**	**100**	**6.4**	**34.5**	**208**	**5**	**7**	**3.4**
	F-2730	TS-9-S	100	5.8	47	176	3	12	6.2

F_s, sampling frequency; depth, hypocentral depth; D_{epi}, epicentral distance; $PGA_{downhole}$, PGA at the downhole station; $PGA_{surface}$, PGA at surface station; F_c, central frequency. Bold indicates the earthquakes used and shown in this article.

peak accelerations at the surface are illustrated in Figure 1 and provided in Table 1. The site-response computations were performed on five input motions at KSRH10 (TS-0-K, TS-1-K, TS-2-K, TS-4-K, and TS-9-K) and four at Sendai (TS-1-S, TS-2-S, TS-5-S, and TS-8-S), numbered from the strongest to the weakest. Only the results of these input motions are shown in this article. Nevertheless, we provide the information for all available input motions as they may be used for future validation exercises.

The PGA and the frequency content of a recording are two relevant parameters of the input motion for describing the expected degree of NL soil behavior (Assimaki and Li, 2012; Régnier, Cadet, *et al.*, 2016). The input motions for KSRH10 and Sendai sites were selected with three different PGA levels (at the downhole sensor), respectively. The PGA was calculated on the acceleration time histories of the geometrical mean of the east–west (EW) and north–south (NS) components, filtered between 0.1 and 40 Hz. In each group of PGA level, we quantified the frequency content using the central frequency following equation (1) (statistical moments order 2 and 0; Sato *et al.*, 1996) but the values were not significantly variable from one event to another. We therefore also considered several magnitudes and epicentral distance couples.

$$F_c = \sqrt{\frac{\int f^2 A^2(f)\,df}{\int A^2(f)\,df}}, \qquad (1)$$

in which F_c is the central frequency, f is the frequency, and $A(f)$ is the amplitude of the Fourier spectrum of the accelerogram. The resulting values exhibit a significant but rather erratic variability, without an obvious link to magnitude, epicentral distance, or depth. We therefore also considered several magnitudes and epicentral distances couples.

The selected events for the KSRH10 and Sendai sites are listed in Table 1 along with their main characteristics (M_w, depth, epicentral distance, PGA at the downhole and at the surface, and central frequency at the downhole recording). The frequency sampling at KiK-net is between 100 and 200 Hz depending on the event and at Sendai it is 100 Hz. For KSRH10, four input motions with PGA at the downhole sensor higher than 50 cm/s² were available and selected, whereas only three were available in Sendai site.

Signal-to-Noise Ratio

We checked the quality of all events by assuring that their signal-to-noise ratios (SNRs) were high enough (SNR > 3) over a broad frequency spectrum of 0.1–50 Hz, for all three components.

Orientation of the Surface to Downhole Sensors

We checked that both surface and downhole sensors are oriented in a similar way. We rotated anticlockwise the surface horizontal components with a 1° azimuth increment, starting from the original EW orientation, and we calculated the correlation coefficient with the downhole EW component. Both signals were filtered between 0.1 and 1 Hz. The angle characterized by the maximum correlation would approximately correspond to the angle between the two EW components of the surface and the downhole sensors. Correlations are maximum without rotation angles at Sendai and at KSRH10. It is close to the calculations performed by Kosaka (PARI) who found a deviation of 7.2° counterclockwise at Sendai also using long-period motions correlations. Results are approximately similar to the results of Maeda *et al.* (2005) who found a deviation of 2.2° clockwise at

KSRH10. These values suggest that both surface and down-hole EW components are mostly oriented parallel one to each other, and even if slight deviations of the order of 7° may occur, it would not impact significantly the soil response functions.

Verification of the Verticality of the Incident Waves

Basic assumptions are made in most 1D simulation codes when propagating a wave through a soil column. One of them is that the wavefield consists of vertically propagating plane *S* waves; hence, the input wave motion at the bottom of the soil column is fully represented by the horizontal components and all vertically propagated toward the surface. However, except for teleseismic (long-distance) events, the seismic waves from local or regional events are very likely to have not only nonvertical incidence (unless located directly underneath the sensor) but also multiple incidence because of crustal scattering.

Recording a nonvertical incident wavefield implies that the total seismic energy is distributed all over the three components. Therefore, if only one component of the 3D wavefield is used to represent one type of seismic phase (in our case, the shear wave), then it is highly likely that the wave energy, PGA, and strains are underestimated at the downhole sensor. However, depending on the thickness of the soil column and its characteristics, the upward propagating body waves become increasingly vertical toward the surface and therefore the surface sensor will give a more complete representation of the entire shear-wave energy on the horizontal components. Therefore, two consequences can result from this observation.

1. Although both the numerical and empirical input motions are the same, their transfer functions would be different, because the energy on the horizontal components at the surface will be greater in the empirical observation at KSRH10 than the numerical simulations.
2. The actual seismic loading at the base of the soil column is underestimated compared to what the soil experiences, and therefore its possible NL behavior can also be underestimated.

To determine the direction of propagation of an input motion, we used a polarization analysis based on the three-component covariance matrix. In Ⓔ Figure S1, we can see that at KSRH10, the polarization analysis indicates that the waves mostly propagate with a vertical incidence (low-incidence angle). At Sendai, except for two recordings, the waves are not linearly polarized; therefore, the calculations of the direction of propagations are not relevant.

From *In Situ* and Laboratory Data to Input Parameters

Identifying the most relevant parameters to be used for simulating the NL wave propagation process in a soil deposit

was one of the main challenges tackled during the verification exercise. For the elastic and viscoelastic properties, V_S, V_P, density, and low-strain attenuation profiles were used. For the NL soil properties, the modulus reduction and damping with shear-strain curves, with the soil shear strength profile were found to be the key parameters to significantly reduce the code-to-code site-response variability (Régnier, Bonilla *et al.*, 2016). For more complicated NL models such as Hujeux (Aubry *et al.*, 1982), more laboratory measurements are required to define its model parameters; however, some of these parameters could be defined using well-known soil mechanics correlations.

The challenge for the validation phase was to determine the value of those parameters for a real site. The specifications for the laboratory and *in situ* measurements were defined in accordance with the prescriptions coming from the organization team with a few associated geotechnical experts and the participating teams, and bounded by the available budget and the measurement capacity of the local company performing the measurements, together with a few logistical issues linked with the exact location of the vertical array and the surrounding environment.

Site Investigation

Measurements Performed.   To obtain the linear and NL soil parameters, *in situ* measurements and multiple laboratory measurements were conducted on disturbed and undisturbed soil samples.

The *in situ* measurements were subcontracted to Oyo company and consisted in (1) boring investigation to determine soil stratigraphy and to perform the soil sampling. The diameter of the borehole was 116 mm up to a depth where triple-tube samplings were used (for sandy soil or relatively stiff clayey soil) then 86 mm; (2) undisturbed soil samples (80 cm long) were collected using the thin-wall sampler for the soft clay soil and using the tripled-tube samplers for the sand and stiffer clayey soil; (3) standard penetration tests; (4) *PS* logging by suspension method for KSRH10 and downhole method for Sendai; and (5) multiple multichannel analysis of surface waves (MASW) at the investigated sites to characterize the spatial variability of the underground structure at shallow depth, together with single-point ambient vibrations recordings. The laboratory soil tests were conducted on disturbed and undisturbed soil samples. The tests on disturbed samples enabled us to determine physical characteristics such as particle size distribution, liquid, and Atterberg limits. The tests on undisturbed soil samples aim at defining the density and to perform a wide range of laboratory tests such as undrained and drained triaxial compressional test, oedometer tests by incremental loading, cyclic undrained and drained triaxial compression test (undrained for investigating the liquefaction potential) and, for rock samples, unconfined compressional tests. The methods used to perform the laboratory tests are defined by Japanese normative specifications.
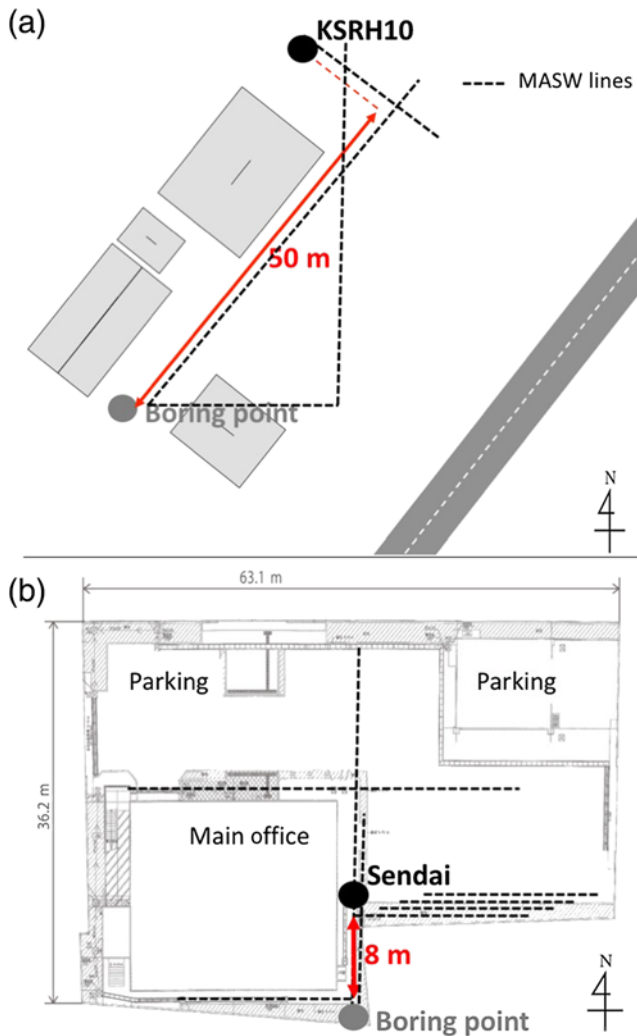
Figure 3. Location of the vertical arrays with respect to the borings for soil parameter measurements and multichannel analysis of surface waves (MASW) lines (a) for KSRH10 site and (b) for Sendai site. The color version of this figure is available only in the electronic edition.

For each borehole, the number of undisturbed soil samples was defined according to the expected soil stratigraphy (on the basis of pre-existing KiK-net or PARI information), to ensure at least one sample in each homogeneous soil unit.

The number and location of soil samples are specified in Table 2 together with the downhole sensor depth and the

maximum depth of the complementary drillings. The details of the locations of the laboratory measurements are available in Ⓔ Figures S2–S5. Figure 3 shows the locations of the boreholes having the accelerometers with respect to the boreholes performed for the laboratory measurements and the MASW lines.

Uncertainties of Soil Parameter Measurements. Because of the inherent variability of the soil and the systems errors in the measurements and sampling methods, a nonnegligible level of uncertainty remains in the soil parameters measured through the laboratory tests. Repeatability of the soil samples and laboratory measurements is a possible way to ensure a reliable definition of the soil parameters. This approach was not applied in this exercise due to budget constraints. We therefore carefully analyzed the data and compared with literature data for similar types of soil.

To minimize the impact of soil spatial variability, the new boreholes were performed as close as possible to the instrumented ones. MASW lines performed between the two boreholes indicate a low spatial variability of the soil parameters for KSRH10 and Sendai (while they did indicate a significant variability at shallow depth for the third site in Onahama, which was also one of the reasons to drop this site for the validation exercise).

Interpretation of the Laboratory and *In Situ* Data

Elastic and Viscoelastic Properties. For the elastic properties, several methods were used to determine the soil parameters. We preferred the methods that provide the direct *in situ* evaluation of the soil properties, yet we did compare the results with alternative techniques characterized by indirect measurements. We used the *PS* logging to obtain the V_S profile and then we used the earthquake recordings to adjust it. As shown in Figure 4, the V_S profile was adjusted to improve the fit between the fundamental resonance frequency recorded and predicted for the KSRH10 and Sendai sites.

For KSRH10, the initial V_S profile was based on the *PS*-logging investigation down to 50 m depth; beyond this depth, we considered the values of the V_S coming from the KiK-net database, where the *PS*-logging method was also used. In this project, it was decided to adjust the linear transfer function from Thomson–Haskell predictions to the instrumental observations of surface–borehole spectral ratios, to ensure that the discrepancies between the prediction and the observations during the benchmark were associated with NL soil behavior, and not to other causes.

To adjust the numerical linear transfer function to the observation, we modify mostly the V_S profile coming from KiK-net for which no information was available on the measurement.

For Sendai, to improve the fit between the weak-motion site response calculated

Table 2
Geological Characteristics of the Two Selected Sites with Locations of the Undisturbed Soil Samples

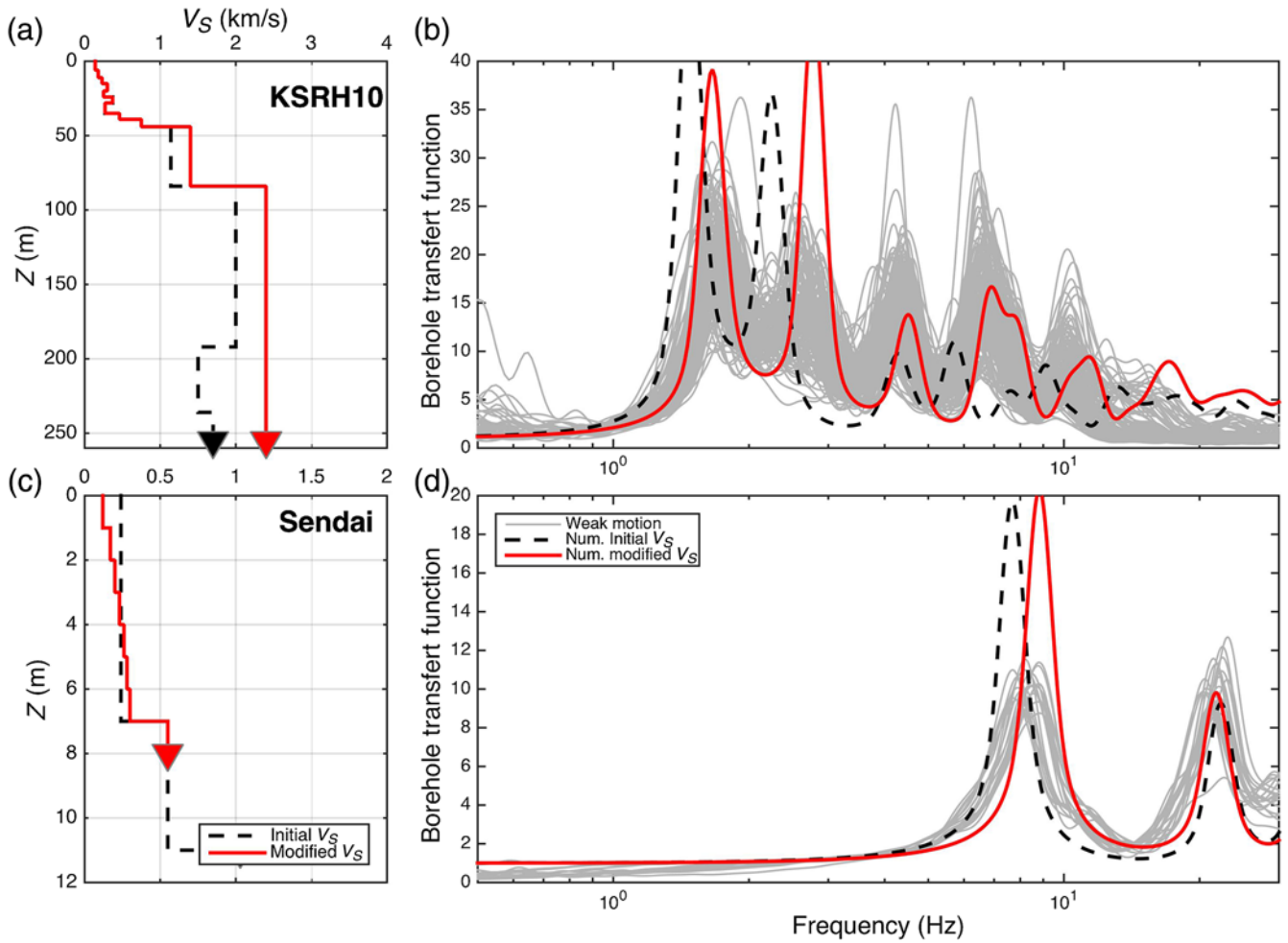| Site	Downhole Sensor Depth (m)	Maximum Complementary Drilling Depth (m)	Type of Soil	Number of Samples (Location)
Sendai	8	10	Sand	2 (3.3 and 5.4 m)
KSRH10	250	50	Sand/clay	6 (3.5, 7.5, 14.5, 22.5, 29.7, and 34 m)

Figure 4. Modification of the V_S profile with improvements of the surface to within spectral ratio with 1D linear site-response analysis. (a,c) The dashed lines represent the initial V_S profile and the plain line of the final V_S profile at KSRH10 and Sendai. (b,d) The dashed lines represent the 1D linear numerical surface to within spectral ratio with the initial V_S profile and the plain line with the final. The color version of this figure is available only in the electronic edition.

with linear site-response analysis and computed from weak motions, a gradient type V_S profile (equation 2) was chosen:

$$V_S = V_{S_1} + (V_{S_2} - V_S) \left[\frac{z - Z_1}{Z_2 - Z_1} \right]^{\alpha}, \qquad (2)$$

in which $V_{S_1} = 140$ m/s, $V_{S_2} = 460$ m/s, and $\alpha = 0.7$, Z_1 is the depth at which it begins the gradient (0 m) and Z_2 (7 m) is the depth where it finishes.

The Poisson coefficient (ν) was computed using the *PS* logging and rounded. To ensure consistency between the values of V_S, V_P, and ν, the V_P parameter was obtained from V_S and the rounded ν. The density was obtained from the undisturbed soil sample and the low-strain attenuation was deduced from the undrained cyclic triaxial test, and, when not available, using the rule of thumb ($Q_S = V_S/10$) (Olsen *et al.*, 2003).

Nonlinear Soil Properties. The initial plan was to use only the measured NL parameters, that is, the degradation curves

measured in the lab. It had however to be modified to ensure a better fit to the strong-motion data: the NL soil properties were actually updated during the iterations of calculations, so that three sets of NL soil parameters were used. The first one (called soil column 1 [SC1]) came simply from the use of NL degradation parameters defined in the literature, and anchored to elastic soil properties. Here, the Darendeli formulation was adopted (Darendeli, 2001). The second (SC2) and third (SC3) parameter sets are directly based on interpretations of the laboratory data. One objective of the benchmark was to focus on routine practice with relatively simple models. Furthermore, the participants were also free to build/ use their own soil model based on the raw experimental laboratory test. Yet, additional NL soil parameters could have been tested as well, such as models that could handle both low and high strain as detailed in Yee *et al.* (2013) and Groholski *et al.* (2016).

Darendeli formulation (Darendeli, 2001) was used to define the G/G_{max} and damping ratio curves as a function of shear strain for SC1. To compute such values, the knowledge

of the confining effective stress (σ'), the overconsolidation ratio, the plasticity index, and the damping ratio at low strains (D_{min}) was required.

The SC2 NL curves were constructed from the cyclic triaxial compression test results. We normalized the Young's modulus decay curves (from the lab fifth cycle of loading) by the low-strain Young's modulus (E_0). E_0 is the value of the hyperbolic model (Hardin and Drnevich, 1972) that mimics the lab results at 0.0001%. We assimilate this E/E_0 decay curve to the shear modulus G/G_{max} decay curve, with G_{max} associated with the *in situ* velocity measurements, that is, $G_{max} = \rho V_{\hat{S}}^2$. The shear strain that used γ was considered equal to 3/2 of the axial strain directly measured during the triaxial test. Indeed, the shear strain is the difference between the axial and radial strain $\gamma = \varepsilon_a - \varepsilon_r$. During the cyclic triaxial test under undrained conditions, volumetric changes are zero: therefore, the volumetric strain is null $\varepsilon_v = \varepsilon_a + 2\varepsilon_r = 0$. From the previous two equations, we can deduce that $\gamma = 3/2\varepsilon_a$ (Vucetic and Dobry, 1988).

The elastic shear modulus values from the laboratory tests (G_{max}^{lab}) are generally underestimated compared to the *in situ* measurements (G_{max}^{insitu}), especially for cyclic triaxial tests (indeed cyclic triaxial tests are not reliable at low strain, below 10^{-4}%). Tatsuoka *et al.* (1995) showed that this could be due to sample disturbance where stronger differences are observed depending on the type of shear-strain measurement. When local measurements of shear strain are performed using internal gauges (inside the soil sample) compared to external measurements (usual measurement), the discrepancies are much smaller.

When normalizing the shear modulus curve to obtain the G/G_{max}^{lab} curve, the G_{max}^{lab} should be corrected. The coefficient of correction to be applied is not well defined but lies between 1.2 and 4 (F. Lopez-Caballero, personal comm., 2015). A correction procedure was set up in this study to partially correct this value (Noguera, 2016). The procedure accounts for the measurement errors when using external measurements instead of local ones but does not consider error due to soil sample disturbance. Considering that this procedure was defined during the project (after the calculations on Sendai site), it was only applied to the laboratory data for KSRH10 site.

The above-mentioned procedure consists of three steps: (1) to find the maximum shear modulus (G_{max}^{lab}) by fitting the logarithmic equation (equation 3) proposed by Nakagawa and Soga (1995), as measured for intermediate strain values; (2) to compare the G_{max}^{lab} values found with those from other tests or in the literature, this value should not be more than twice the G_{max}^{lab} as it only accounts for the external-to-local strain measurement error; and (3) to normalize the fitted decay curve and multiply it by G_{max}^{insitu} value:

$$G/G_{max}(\gamma) = \frac{1}{1 + \alpha|\gamma|^\beta}. \tag{3}$$

The SC3 model was built using the hyperbolic model (equations 4 and 5) constrained by the G_{max}^{insitu} and the shear

strength (τ_{max}). The latter was estimated from the depth and the cohesion and friction properties according to equation (6). This formula was derived from the Mohr circle for simple shear test (Hardin and Drnevich, 1972). The shear strength used here is the maximum shear stress not the maximum lateral shear stress, with the effective cohesion (c') and the friction angle (φ') coming from the monotonic compressional test and the coefficient of soil at rest (K_0) coming from Jaky's formula ($1 - \sin(\varphi')$; Jaky, 1944).

$$G/G_{max}(\gamma) = \frac{1}{1 + \gamma/\gamma_{ref}} \tag{4}$$

$$\gamma_{ref} = \frac{\tau_{max}}{G_{max}} \tag{5}$$

$$\tau_{max} = \frac{(1 + K_0)\sigma'_0 \sin \varphi'}{2} + c' \cos \varphi'. \tag{6}$$

The comparison of the NL curves is illustrated in Figures 5 and 6 for KSRH10 and Sendai, respectively. For both sites, the SC1 NL curves have generally more shear modulus reduction for the same shear strain than those coming from the laboratory data, even after the correction procedure was applied.

The mechanical properties of the KSRH10 and Sendai sites are synthesized in Tables 3 and 4, respectively. For KSRH10, three sets of NL parameters were tested, whereas only the first 2 were tested on Sendai site. From *in situ* surveys, the water table is located, respectively, at 2.4 and 1.45 m below the ground surface for KSRH10 and Sendai site.

Validation Results

Calculations Performed

Two iterations of calculations were performed at KSRH10 and Sendai sites. The first iteration was completely blind (i.e., only input motions were given to the participants), while during the second one the surface motions were also provided. Three soil column models were tested for KSRH10 and two for Sendai. Both EW and NS components were used. In addition, the rotated horizontal motion corresponding to the maximum peak acceleration was considered; simulations using this rotated horizontal component were performed. However, the results were not significantly different from those obtained using the EW or NS components. Thus, these computations are not shown here.

All the participating teams were asked to provide the acceleration time histories and the stress–strain curves at different depths in the soil columns. For KSRH10, the acceleration time histories were computed at various ground levels from the surface (GL): GL-0, GL-6, GL-11, GL-15, GL-20, GL-24, GL-28, GL-25, GL-39, GL-44, GL-84, and GL-255 m depth, corresponding to the main soil layer interfaces. The stress–strain curves were computed at GL-3, GL-8.5, GL-13, GL-17.5, GL-22, GL-26, GL-31.5, GL-37, GL-41.5, GL-64, and GL-169.5 m, which correspond to the
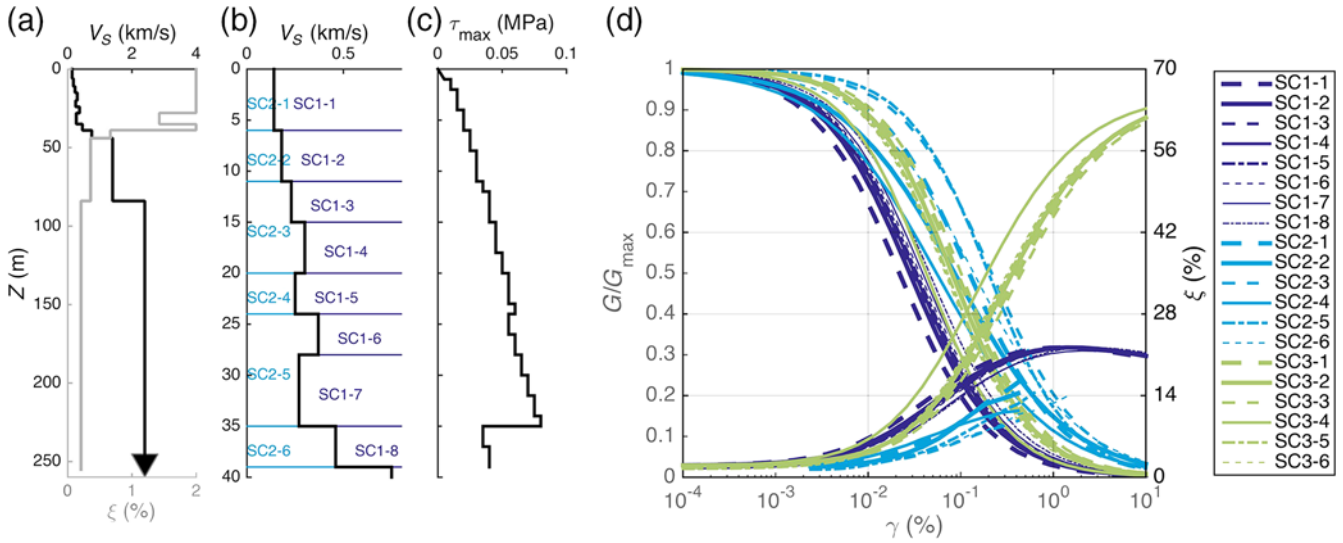
Figure 5. Input parameters for the numerical simulations at KSRH10 site. (a) V_S and elastic attenuation (ξ_0) profiles. The graph (b) illustrates, in the soil layer down to the depth where nonlinear (NL) soil behavior is defined, the location of the shear modulus decay and attenuation ($G/G_{max}(Y)$ and $\xi(Y)$) curves for the soil column 1 (SC1), for SC2, and for SC3 shown in (d). For SC3, the locations of G/G_{max} and damping curves in the soil layers are similar as for SC2. (c) The shear strength profile down to 40 m. The color version of this figure is available only in the electronic edition.
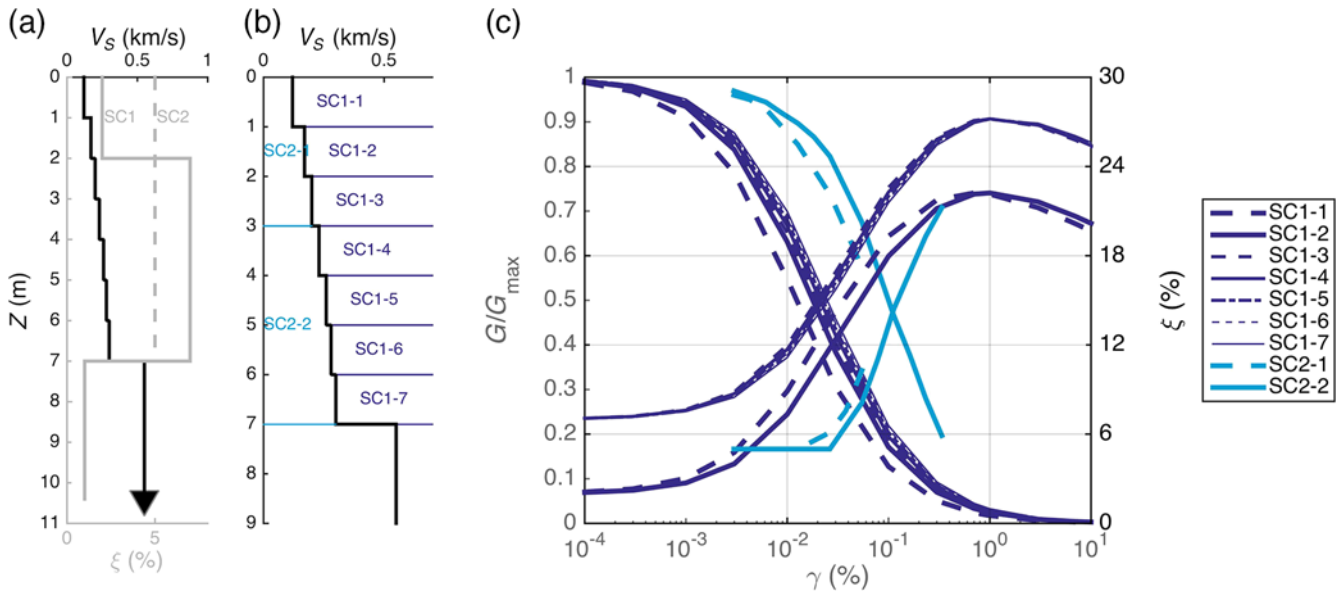

Figure 6. Input parameters for the numerical simulations at Sendai site. (a) V_S and elastic attenuation (ξ_0) profiles. The graph (b) illustrates, in the soil layer down to the depth where NL soil behavior is defined, the location of the shear modulus decay and attenuation ($G/G_{max}(Y)$ and $\xi(Y)$) curves for SC1 and for SC2 shown in (c). The color version of this figure is available only in the electronic edition.

middle of the soil layers. For Sendai, the acceleration time history was provided from GL-0 to GL-8 m every 1 m, whereas the stress–strain curves were provided from GL-0.5 to GL-7.5 m also every 1 m.

Analysis of Results

This article focuses on the analysis of the whole dataset returned from each team to estimate the level of uncertainties associated with NL modeling.

Several sources of uncertainties are involved in 1D NL site-response analyses. On the one hand, there are epistemic uncertainties coming from the main assumption of the method (1D, vertical propagation of *SH* waves), soil parameters measurements (which should however impact all the predictions in the same way), the numerical model, and the users. On the other hand, there are random uncertainties coming from the input motions, which are influenced by both the seismic sources and soil heterogeneities. We did use

Table 3
Soil Properties from the KSRH10 Site

Z (m)	V_S (m/s)	V_P (m/s)	ρ (kg/m³)	Q_S	ξ	Set of G/G_{max} and Damping Curves	τ_{max} (kPa)
6	140	1520	1800	25	0.02	SC1-1, SC2-1, SC3-1	Calculated every 1 m
11	180	1650	1800	25	0.02	SC1-2, SC2-2, SC3-2	(according to equation 6)
15	230	1650	1500	25	0.02	SC1-3, SC2-3, SC3-3	
20	300	1650	1500	25	0.02	SC1-4, SC2-3, SC3-3	
24	250	1650	1600	25	0.02	SC1-5, SC2-4, SC3-4	
28	370	1650	1600	25	0.02	SC1-6, SC2-5, SC3-5	
35	270	1650	1800	35	0.0142	SC1-7, SC2-5, SC3-5	
39	460	1650	1800	25	0.02	SC1-8, SC2-6, SC3-6	
44	750	1800	2500	75	0.0066	Linear	
84	1400	3400	2500	140	0.0035	Linear	
255	2400	5900	2500	240	0.0020	Linear	

Z, depth of the soil layer; V_S, shear-wave velocity of the soil layer; ρ, density of the soil layer; Q_S, elastic damping ratio; ξ, elastic attenuation; τ_{max}, shear strength; SC, soil column.

Table 4
Soil Properties from the Sendai Site

Z (m)	V_S (m/s)	V_P (m/s)	ρ (kg/m³)	Q_S	ξ	Set of G/G_{max} and Damping Curves
1	120	610	1850	25	0.02	SC1-1, SC2-1
2	170	870	1850	25	0.02	SC1-2, SC2-1
3	200	1040	1850	7.14	0.07	SC1-3, SC2-1
4	230	1180	1890	7.14	0.07	SC1-4, SC2-2
5	260	1300	1890	7.14	0.07	SC1-5, SC2-2
6	280	1420	1890	7.14	0.07	SC1-6, SC2-2
7	300	1530	1890	7.14	0.07	SC1-7, SC2-2
10.4	550	2800	2480	50	0.01	Linear

several input motions; however, the number of them is not sufficient to take into consideration all the random uncertainty ranges. The previous verification phase (Régnier, Bonilla, *et al.*, 2016) provided an estimation of the code-to-code variability linked to the numerical method such as numerical integration schemes, implementations of damping, constitutive models, and users practice. Conversely, the validation phase involves comparison with observations and therefore calculations of residuals. We assume that the residual can be described as a random variable with normal distribution center around a nonnecessarily zero mean (models may overpredict or underpredict), with associated standard deviation.

We calculated and compared the misfit between the observations and the computations with the code-to-code variability of the surface-response spectra averaged over a period bandwidth close to the resonance period of the site, namely $[0.7f_0, 1.3f_0]$. The misfit reflects the total uncertainty (epistemic and random) of the results, whereas the code-to-code variability illustrates the part of the epistemic uncertainty associated with the choice of a numerical model. Then, this code-to-code variability is compared to the previous results obtained in the verification exercise done on simplified soil profiles.

An additional analysis was performed to consider the soil column choice in the whole uncertainty assessment.

The misfit was calculated for each input motion and all soil columns (SC1, SC2, and SC3) together.

Finally, the results at each site were analyzed, through the distribution of the computed transfer functions and response spectra, using the 25th and 75th percentiles of all computations, and comparing them with the observations. Concomitantly, we also computed and analyzed the distribution of residuals on response spectra to quantify the discrepancy and identify when the observations were underestimated and overestimated. This operation was carried out at each oscillator period.

Code-to-Code Variability versus Misfit

A first analysis of the code-to-code variability (hereafter, σ_{c2c}) relative to the variability of the residuals between the recording and the simulations (hereafter, Misfit) is provided.

To quantify the Misfit, we calculate the root mean square distance (rmsd) between each prediction with the observation of the response spectra as proposed in equation (7) and averaging it (geometric mean) over the periods of interest of the site:

$$\text{Misfit} = \left(\prod_{i=1}^{n} \text{rmsd}_{\text{obs-num}}(T_i) \right)^{\frac{1}{n}},$$

in which

$$\text{rmsd}_{\text{obs-num}}(T_i) = \sqrt{\frac{1}{N_c - 1} \sum_{j=1}^{N_c} [\text{LSA}_{\text{obs}}(T_i) - \text{LSA}_{\text{num,j}}(T_i)]^2},$$
(7)

in which n is the number period sample between $T_1 = 1/(1.3f_0)$ and $T_2 = 1/(0.7f_0)$ of the discrete response spectral acceleration (SA); $\overline{\text{LSA}_{\text{obs}}}$ is the logarithmic (to base 10) transformation of the observed SA; $\text{LSA}_{\text{num,j}}$ is the logarithmic transformation of jth surface predicted SA.

The code-to-code variability (σ_{c2c}) is the standard deviation of the predictions as defined in equation (8), averaged over the same period range as before:
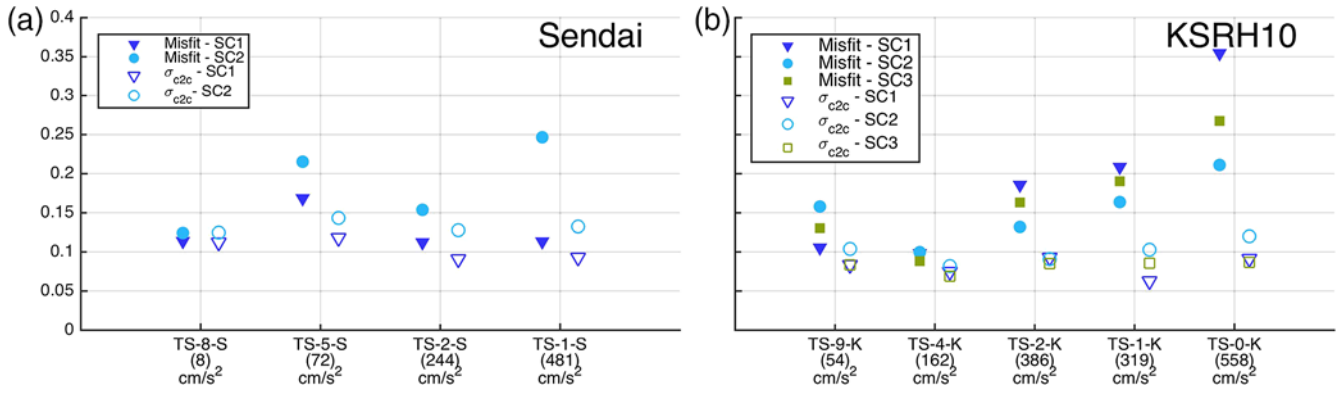
Figure 7. Distance between the recorded and computed surface pseudoresponse spectra (Misfit) compare to the code-to-code variability (σ_{c2c}) at Sendai site, using the SC1 and SC2 for the input motions TS-1-S, TS-3-S, TS-5-S, and TS-9-S and at KSRH10 using SC1, SC2, and SC3 for the input motions TS-0-K, TS-1-K, TS-2-K, TS-4-K, and TS-8-K. The color version of this figure is available only in the electronic edition.

$$\sigma_{c2c} = \left(\prod_{i=1}^{n} \mathrm{rmsd}_{\mathrm{num}}(T_i) \right)^{\frac{1}{n}},$$

in which

$$\mathrm{rmsd}_{\mathrm{num}}(T_i) = \sqrt{\frac{1}{N_c - 1} \sum_{j=1}^{N_c} [\overline{\mathrm{LSA}_{\mathrm{num}}}(T_i) - \mathrm{LSA}_{\mathrm{num},j}(T_i)]^2},$$

(8)

in which n is the number period sample between $T_1 = 1/(1.3 f_0)$ and $T_2 = 1/(0.7 f_0)$ of the discrete response SA; $\overline{\mathrm{LSA}_{\mathrm{num}}}$ is the logarithmic (to base 10) transformation of the mean predicted SA; $\mathrm{LSA}_{\mathrm{num},j}$ is the logarithmic transformation of jth surface predicted SA.

The considered period range spans an interval of $\pm 30\%$ around the fundamental resonance frequency of the sites (f_0) (8.2 and 1.7 Hz for Sendai and KSRH10, respectively). For Sendai, it corresponds to the frequency range from 5.47 to 10.66 Hz, equivalent to periods between [0.09–0.18] s. For KSRH10, the adopted frequency interval is [1.19–2.21] Hz, corresponding to periods between [0.45-0.84] s. The periods are log-scale sampled.

Figure 7 compares the Misfit on the EW component for Sendai and the NS component for KSRH10 (filled markers), with the code-to-code distance σ_{c2c} (empty markers). The simulated component (EW or NS) with less discrepancy with the observations was chosen.

For Sendai, we compared the results for four input motions (1, 2, 5, and 8, for which all calculations were performed) and for the two soil models SC1 and SC2. For KSRH10, it is illustrated for five input motions (0, 1, 2, 4, and 9) and for the three soil models (SC1, SC2, and SC3). As expected, the misfit is systematically higher than the code-to-code variability regardless of the input motion or site considered: actually they could be equal only if the predictions are unbiased in average, that is, if $\overline{\mathrm{LSA}_{\mathrm{num}}}(T_i) = \mathrm{LSA}_{\mathrm{obs}}(T_i)$ for every period T_i—which actually never happens. It is also worth observing that the code-to-code variability is quite the same regardless of the input motion, the site, or the soil columns

considered and has a minimum value of 0.06 and a maximum value of 0.15. This suggests that for 1D seismic response analyses, the epistemic uncertainty related to the choice of the numerical method and soil constitutive model should be considered between 0.06 and 0.15 (in log10 scale). Those values, when compared to well-known uncertainty estimation such as in GMPEs, lying between 0.15 and 0.35 (Strasser *et al.*, 2009), are significant and should be taken into account for seismic hazard assessment. One must also keep in mind that such values correspond to a relatively narrow frequency range, and may not be representative of the variabilities in other frequency ranges (as may be seen on Fig. 8).

The misfits are generally lower for the weakest input motions regardless of the site considered, although slightly more pronounced for the KSHR10 site. This is expected because the predictions are closer to one another and also closer to the observations when the soil response is mainly in the linear range. An exception is observed at Sendai for TS-5-S, for which the misfit is larger for a moderate PGA. The misfits between observations and simulations are found significantly lower at Sendai than at KSRH10. They lie between 0.1 and 0.25 at Sendai, whereas they are in the interval between 0.08 and 0.35 for KSRH10.

Regarding the soil column, the SC1 model provided closer results to the observation compared to SC2 at Sendai. Conversely, for KSRH10, SC2 and SC3 models led to lower misfit values.

Comparison between Verification and Validation Epistemic Variability

In this section, the variability of the predictions performed during (1) the verification phase on canonical cases (three profiles P1, P2, and P3 with fundamental resonance frequencies of 3.75, 1.16, and 1.58 Hz, respectively) and (2) the validation phase on real sites are compared in terms of standard deviation (log10 unit) of the PGA and SAs at the surface.

We provide the results in Figure 8. For the validation phase, the standard deviation is calculated for KSRH10 for
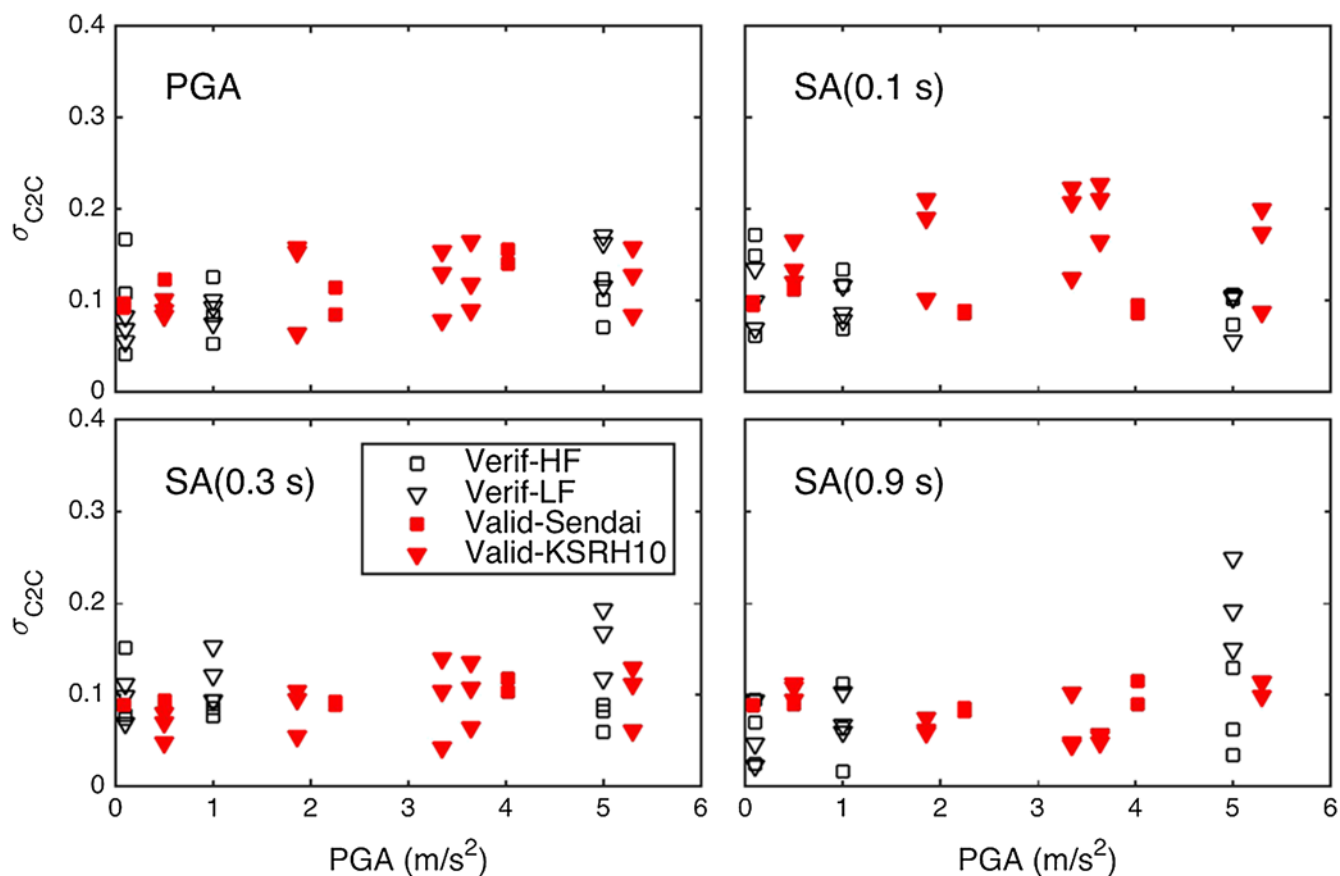
Figure 8.   Standard deviation values of the logarithm results (PGA, spectra acceleration at three periods) for the verification phase on canonical cases and for the validation phase of PRENOLIN project. The color version of this figure is available only in the electronic edition.

the input motions: TS-9-K, TS-4-K, TS2-K, TS-1-K, and TS-0-K. For Sendai, it is calculated for the input motions TS-8-S, TS-5-S, TS-2-S, and TS-1-S. As far as the verification phase is concerned, we considered only the results for the first profile P1, the rigid substratum case, and the NL computations. The numerical results depend on the input motion level and on the frequency content (HF stands for high-frequency input motion and LF for low frequency, see Régnier, Bonilla, *et al.*, 2016); the variability increases with the strain level developed in the soil column. It is therefore higher at high PGA and for the low-frequency content input motion. The LF waveform generates higher strains compared to the HF waveform, for the same PGA level (black empty triangles). This is because the frequency content of the input motion is close to the resonance frequency of the canonical site. Therefore, strong resonance effects are expected. In the validation phase, the variability is generally larger for the stronger input motions except at KSRH10 site for periods above 0.3 s.

Propagation of the Epistemic Uncertainty

We built a logic tree similarly to what is done in probabilistic seismic hazard assessment, to propagate the uncertainty of the numerical simulation and interpretation of the soil data from *in situ* and laboratory measurements to the site-response

and surface-response spectra assessment. For each site (Sendai or KSRH10) and each input motion, this logic tree is composed of two nodes (as shown in Fig. 9). The first node is the soil column (SC1 and SC2 for Sendai; SC1, SC2, and SC3 for KSRH10) and the second node is relative to the team and code couple (from EA-0—team A with his first code—to EZ-1—team Z with his second code). All branches of the tree have the same weight. The uncertainty is quantified by the standard deviation of residuals of the logarithm (log10 unit) of the results (here PGA and response spectra at the surface at three periods 0.1, 0.3, and 1 s) as defined in equation (7).

Table 5 synthesizes the standard deviation of the results obtained in the present PRENOLIN exercise. It might be noted that the rms of residuals (rmsd) are in most cases lower for Sendai compared to KSRH10, except for longer periods (1 s) (period close to the KSRH10 fundamental resonance frequency where the fit is good). The fit is generally better for weak input motions, except at Sendai for the response spectra above 0.1 s.

Comparison of Transfer Function and Response Spectra between Soil Columns

Let us analyze more precisely for each period (frequency) the differences between predictions and observations. The
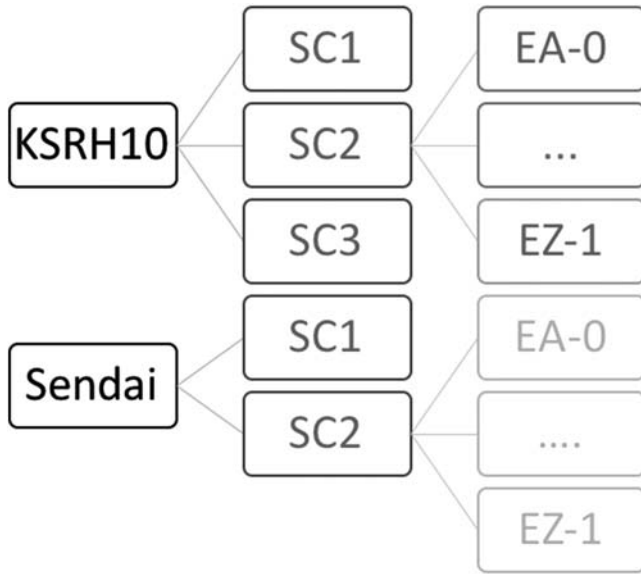
Figure 9. Logic tree for propagation of the epistemic uncertainty.

25th and 75th percentiles of the surface-response spectra and the borehole transfer functions are compared with the observations, for a strong and a weak input motion. We selected TS-1-S and TS-8-S for Sendai and TS-1-K and TS-9-K for KSRH10.

Besides, to quantify the discrepancy between observations and predictions, the average residual (R) per period was calculated according to the following equation:

$$R(T) = \frac{\sum_{j=1}^{N_c}[\mathrm{LSA_{obs}}(T) - \mathrm{LSA_{num},}_{j}(T)]}{N_c}, \qquad (9)$$

in which N_c is the number of computations; $\overline{\mathrm{LSA_{obs}}}$ is the logarithmic (base 10) of the observed SA; and $\mathrm{LSA_{num},}_{j}$ is the logarithmic of the jth surface predicted SA.

Sendai. In Sendai, the results of the computations are closer to the observations when using soil model 1 (SC1), which was defined using literature parameters. In Figure 10, the transfer functions are compared, the fundamental resonance frequency (f_0) of the observations for a weak input motion (TS-8-S) is equal to 8.5 Hz, which is well reproduced

by the numerical computations using either SC1 or SC2 parameters. For the strongest input motion (TS-1-S), f_0 is equal to 7.3 Hz in the observations. In the computations, f_0 is similar when using SC1 parameters while for SC2, f_0 is slightly above (7.8 Hz), indicating a lower level of nonlinearity when using SC2 parameters compared to SC1. Similarly, as illustrated in Figure 11, the surface-response spectrum is well reproduced by the computations using the two soil columns for weak motion (TS-8-S), whereas for the strong motion (TS-1-S) the prediction using SC1 is closer to the observation as compared to SC2. Those observations are also highlighted in Figure 12, and the residuals are close to 0 for the weakest input motion (TS-8-S) for which the two soil columns provide similar estimations. For the strongest input motion, we observe an overestimation below 0.2 s. For TS-5, the underestimation is observed for both soil columns, whereas for TS-1-S and TS-2-S it is mainly observed for SC2. It shows that the discrepancy between the observations and the predictions for TS-5-S does not depend on the soil column characteristics and may be associated with the input motion specificity.

These first results were somehow disappointing because site-specific measurements failed in predicting the observations where generic parameters succeeded. We investigate the source of this discrepancy that could come either from a measurement error or a misinterpretation of the laboratory tests. For Sendai site, there is a large variability between the measured laboratory shear modulus and the *in situ* measurement. The shear modulus from the laboratory measurement is equal to 25 MPa at 3.3 m depth, compared to 100 MPa from the *in situ* measurement of V_S (230 m/s) and density values (1890 kg/m³). This observation suggests that the correction of the laboratory data should have been applied to Sendai site as well.

The procedure to correct the G/G_{max} curves, as indicated previously, has not been applied to the laboratory at Sendai data before the calculations, but it was performed *a posteriori*.

The comparison of the G/G_{max} curves of SC2 model with G/G_{max} curves from laboratory data did not indicate modifications that could explain the large misfit for SC2 model. We recall that this procedure is supposed to correct only for measurement errors between external and local shear-strain devices. Considering the large uncertainty that can lie in the value of the G_{max}^{lab}, it is highly recommended that low-strain measurements such as resonant column or bender element should be used in addition to the cyclic triaxial test to define these parameters.

KSRH10. KSRH10 is a deep sedimentary site with downhole station at GL-255 m. We can observe that the site response (Fourier transfer function) is variable depending on the component of motion and hard to reproduce above the fundamental resonance peak.

Table 5

Standard Deviation Values of the Residuals of the Logarithm Results for All Teams and Soil Columns (PGA, Spectral Acceleration [SA] at Three Periods)

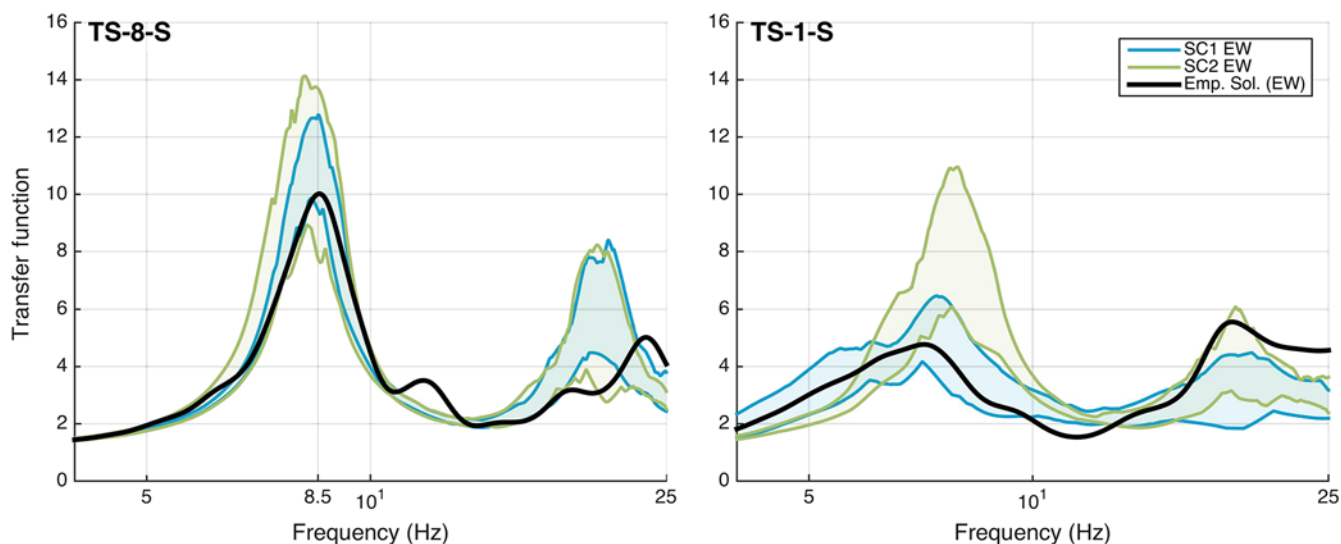|	Sendai				KSRH10				
	TS-1-S	TS-2-S	TS-5-S	TS-8-S	TS-0-K	TS-1-K	TS-2-K	TS-4-K	TS-9
PGA	0.16	0.11	0.11	0.09	0.20	0.17	0.17	0.17	0.10
SA (0.1 s)	0.10	0.10	0.12	0.10	0.23	0.22	0.23	0.20	0.15
SA (0.3 s)	0.11	0.09	0.09	0.09	0.17	0.14	0.15	0.09	0.07
SA (1 s)	0.10	0.08	0.09	0.09	0.11	0.06	0.07	0.07	0.11

Figure 10.   Comparison of the empirical surface to within spectral ratio of the east–west (EW) of the input motion TS-1-S and TS-8-S recorded at Sendai with the envelope of the results represented by the 25th and 75th percentiles of all numerical computations using SC1 or SC2 and only the EW component. The color version of this figure is available only in the electronic edition.
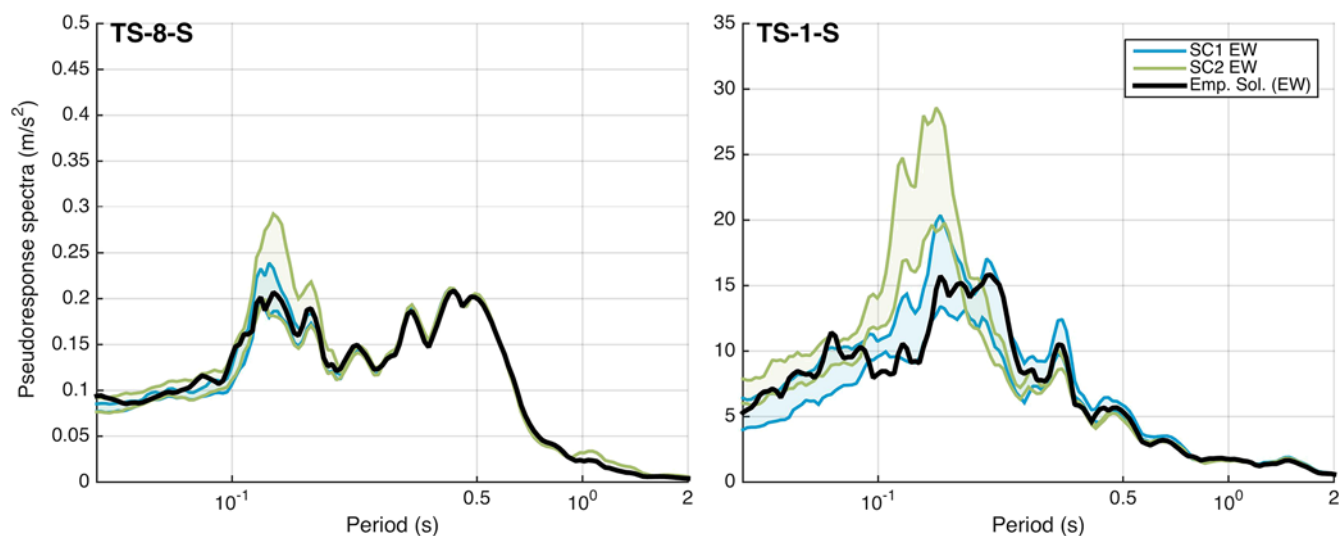

Figure 11.   Comparison of the surface-response spectra of the EW components of the input motion TS-1-S and TS-8-S recorded at Sendai with the envelope represented of the 25th and 75th percentiles of all numerical computations using SC1 and SC2. The color version of this figure is available only in the electronic edition.

In Figure 13, the recorded transfer function is variable from one component of motion to another especially for TS-1-K, where the amplitude of the fundamental resonance peak is equal to 25 for the NS component and only 16 for the EW component. The frequency peaks between 2 and 5 Hz are variable from component to component and event to event and could not be predicted by the 1D assumptions made in this study. For TS-1-K, we observe second and third peaks at 2.3 and 3.3 Hz for the EW component and at 2.7 and 3.2 Hz for the NS component, whereas only one peak is predicted by the numerical simulation with a very high amplitude at 2.7 Hz (high amplitude in the surface Fourier spectrum). For

TS-9-K, we observe three peaks at 2.3, 3.2, and 4 Hz and only one is predicted by numerical simulations at 2.6 Hz. The fourth peak, close to 7 Hz, is more stable from one component to another, but the frequency is slightly overestimated and the corresponding amplitude is underestimated when using SC1.

At high frequencies (above 12 Hz), a deamplification is observed in the empirical borehole transfer function that is not reproduced by the simulations. One possible explanation is the existence of a noticeable soil–structure interaction at the accelerometer sites: this was proposed by Disaster Prevention Research Institute (DPRI) based on their own
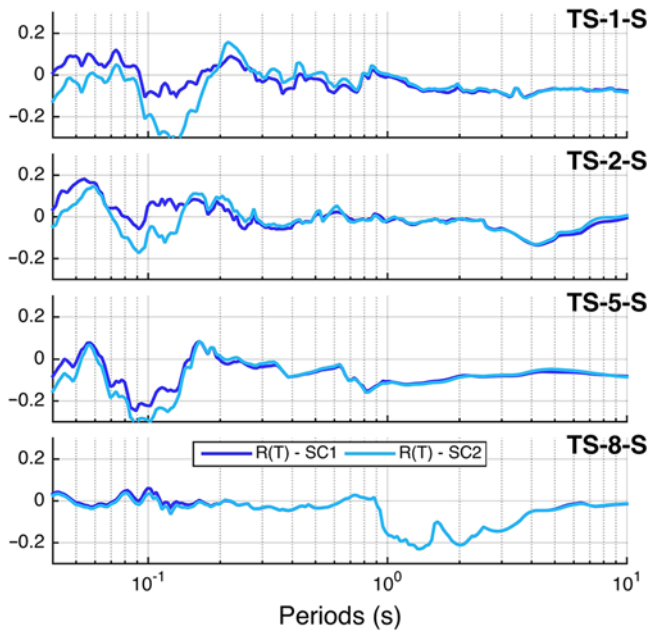
Figure 12. Comparison of the residuals and associated standard deviation of the surface-response spectra of the EW component of the input motion TS-1-S, TS-2-S, TS-5-S, and TS-8-S recorded at Sendai with the envelope represented by the 25th and 75th percentiles of all numerical computations using SC1 and SC2. The color version of this figure is available only in the electronic edition.

experience at several sites and many events. This may also add to some side effects of the low-pass filtering of all KiK-net recordings below 25 Hz. Other possible contributions to this high-frequency reduction may be larger damping (especially for low strains) or scattering from shallow small-size heterogeneities.

For the fundamental resonance peak, we found that SC2 and SC3 coming from *in situ* measurements provided closer results to the observations than SC1 defined using literature

G/G_{max} curves (Darendeli, 2001). As shown in Figure 13, for TS-1-K input motion, the amplitude of the observed resonance frequency peak at 1.4 Hz is around 25 for the NS component, whereas SC1 predicted amplitudes (in the 25th–75th percentiles envelope) between 12 and 14 only.

As seen in Figure 14, the amplitude of the surface-response spectra for the NS component is relatively well reproduced for the weak input motion TS-9-K, but the periods of maximal amplitudes are shifted, creating an overestimation at 0.13 and 0.34 s and an underestimation at 0.18 and 0.25 s. These discrepancies are related to the differences between the observed and computed frequency peaks in the transfer function. For TS-1-K, the amplitude is well reproduced except for SC1, for which the amplitude is significantly underestimated between 0.1 and 0.34 s. The surface-response spectra from the computations using SC2 and SC3 are closer to the observations.

In Figure 15, we observe the residuals of the response spectra; the recordings are underestimated and overestimated especially for periods close to 0.35 s for TS-1-K–TS-9-K input motions. The underestimation amplitude increases with the input motion intensity and is more important for SC1.

At KSRH10, the site-specific measurements on NL properties provide more satisfactory results than the generic curves. The type of soil of KSRH10 was analyzed in detail by one of the participants (G. Lanzo, personal comm., 2015). The NL parameters (normalized shear modulus reduction and damping curves) defined in SC2 and SC3 are less NL compared to classical literature curves even for similar types of soil (i.e., SC1). This observation is consistent with the fact that KSRH10 is composed of volcanic sand, as indicated by site geology. Several authors have shown that volcanic sand exhibits a lower NL behavior as compared to classical sand. For example, laboratory experimental tests conducted by
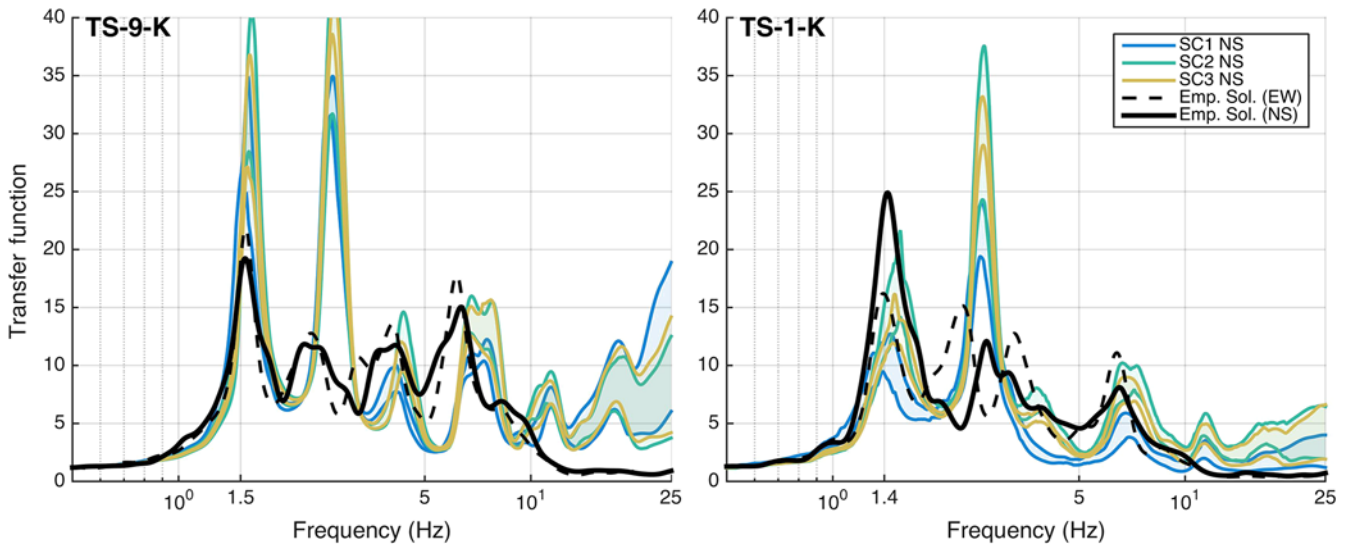

Figure 13. Same as Figure 10. For KSRH10, for the input motions TS-1-K and TS-9-K and using SC1, SC2, and SC3 for the north–south (NS) components. The color version of this figure is available only in the electronic edition.
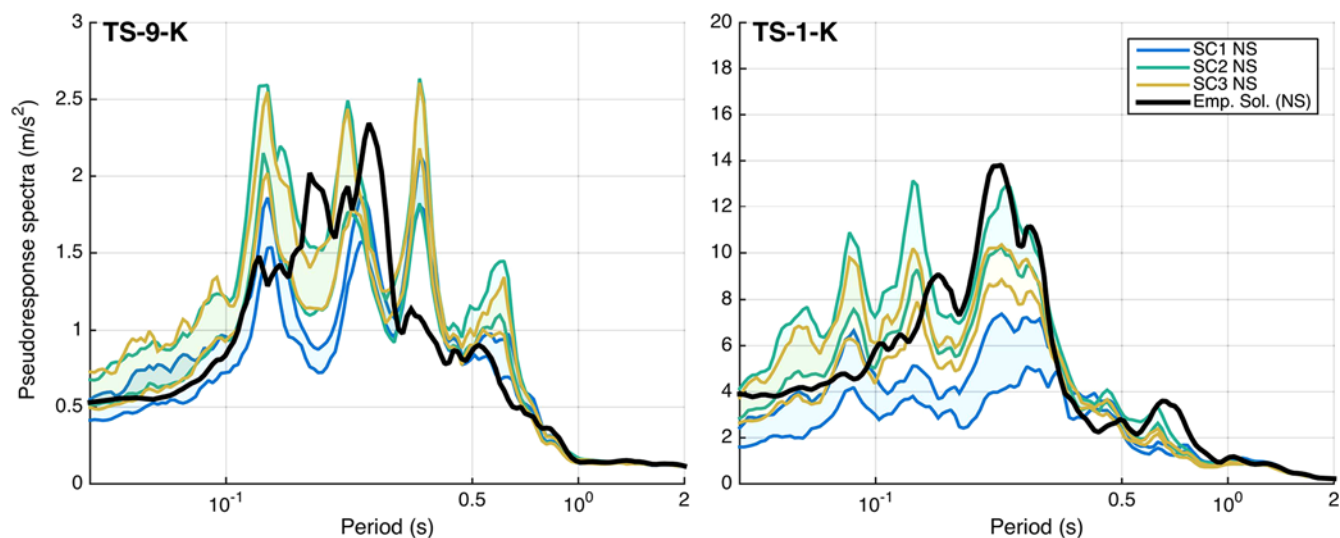
Figure 14.   Same as Figure 11. For KSRH10, for the NS component of the input motions TS-1-K and TS-9-K and using SC1, SC2, and SC3. The color version of this figure is available only in the electronic edition.

Senetakis *et al.* (2013) show that pumice sands exhibit a slower normalized stiffness decay and a lower dissipative behavior than classic gravel curves (Rollins *et al.*, 1998). Similar experimental results have been obtained in Italy on volcanic materials such as Colle Palatino tuff in Rome (Pagliaroli *et al.*, 2014), Naples Pozzolana (Papa *et al.*,

Figure 15.   Same as Figure 12. For KSRH10, for the NS component of the input motions TS-0-K, TS-1-K, TS-2-K, TS-4-K, and TS-9-K and using SC1, SC2, and SC3. The color version of this figure is available only in the electronic edition.

1988), and Orvieto (Central Italy) pyroclastic materials (Verrucci *et al.*, 2015). Thus, Darendeli's curves built on classical sand data could not necessarily reproduce the NL soil behavior at this site. This exercise suggests that one should be careful when using generic soil curves such as Darendeli (2001) or others: the soil nature is important to evaluate their relevancy in site-response analyses.
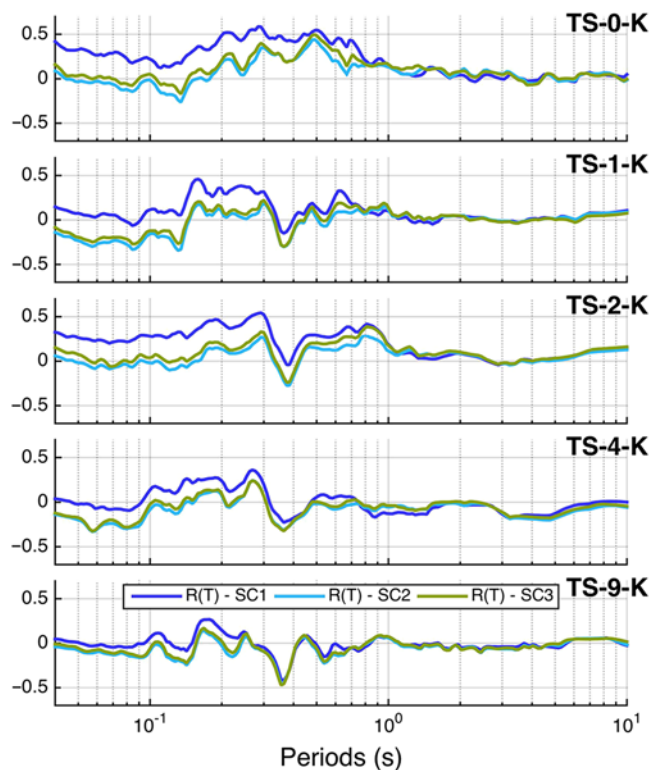
Discussion and Conclusions

Applicability of the Calculations

The computations performed in this benchmark are limited by three main assumptions: (1) 1D wave propagation in horizontally layered media, (2) *SH* waves (only one component of motion) with vertical incidence, and (3) total stress analysis.

1D Structure.   We succeed at reproducing the fundamental resonance frequency at both sites (for weak and strong motions), but the higher modes remain difficult to reproduce. The sites were chosen (over 688 sites in KiK-net) to fulfill specific criteria warranting limited deviations from a 1D site configuration. We observe at KSRH10 site that 1D numerical simulation could not reproduce the observed site response over the whole frequency range, even for weak motions: this is likely to indicate that the site has more complex geometry. The 1D structure assumption is a very strong one which may not be realistic. However, moving forward for more complex geometries requires more detailed site characterization over a broader area, adequate interpretation of the data, and application of 2D and 3D numerical simulations (Dupros *et al.*, 2010; Taborda *et al.*, 2010; Amorosi *et al.*, 2016). Benchmarking NL numerical codes for 2D or 3D geometries is a real challenge, which should start with a carefully designed verification exercise.

Vertically Incident SH Waves. Vertically incident *SH* waves' loading implies two distinct assumptions: (a) vertically incident plane waves and (b) a unidirectional motion in the whole soil column.

The former has been partially tested through a polarization analysis and has been found only partially fulfilled. In any case, there is not enough information from the two-sensor recordings to constrain the complexity of the incident wavefield, which is an unknown mixture of body waves with varying incidence angles and back azimuths, and surface waves with varying back azimuths.

Concerning the second assumption, one participant tested a code with an implementation of a three component (3C) NL constitutive relation. Using a 3C NL constitutive model has been shown to be relevant for strong ground motion prediction with 1D wave propagation models for a large event such as 2011 Tohoku (Santisi d'Avila and Semblat, 2014). The 3D loading path due to the 3C polarization leads to multiaxial stress interaction that reduces soil strength and increases NL effects.

Therefore, for Sendai site, results of 1D analysis performed using 3C codes (1D, 3C) were compared to those obtained by 1D analysis with only one component of motion (1D, 1C, vertical incidence; see Ⓔ Fig. S6). Simulations were carried out with reference to SC2, which is less NL than SC1, and considering the strongest input motion (TS-1-S).

As illustrated in Ⓔ Figure S6, no significant differences between the two results were observed and cannot explain the discrepancy between the predictions and the observation. Additional research is needed to further explore the impact of 3C motions versus 1C assumption.

Total Stress versus Effective Stress Analysis. Even though the exercise was limited to a total stress analysis, and designed accordingly, we thought it useful to benefit from the willingness of some volunteer participants to investigate whether the type of analysis (effective or total stress) or the input soil model used could improve the fit. Two teams (W-0 and H-0) performed those calculations at Sendai site. Ⓔ Figure S7 presents the borehole transfer function computed for SC1 and SC2 sorted according to the type of analysis (total stress analysis, EQL, or effective stress analysis), as compared to the observations for two different input motions (TS-1-S and TS-8-S). We observe that the weak-motion effective stress analysis provides results close to total stress analysis, as expected. This is also the case for the strongest input motion when dealing with SC2 parameters. However, one team (W-0) succeeds in reproducing the observations even with SC2 parameters, when using an effective stress analysis. That team indicated that the NL input data used in this analysis were calibrated directly from the laboratory tests and adjusted to the *in situ* measurements of elastic properties. The G/G_{max} curves obtained were closer to the curves used in SC1 rather than SC2. Therefore, no evidence of efficiency of effective stress analysis compared to total stress analysis is available for Sendai site. Additional research is

required to further evaluate the conditions under which an effective stress analysis is needed.

Equivalent Linear Method. In addition to the time-domain NL site-response analyses limited by these three assumptions, the EQL method was also tested. This approach involves a linear computation, coupled to an iterative process that adjusts, at each iteration, the value of the shear modulus and damping, according to the maximum shear strain calculated at the middle of each soil layer. This method is largely used in earthquake engineering practice since the pioneering work of Schnabel *et al.* (1972). Three teams used an EQL method (J-1 and Z-0). Team J-1 performed all the calculations, whereas Z-0 provided the results of the EQL method only for the weakest input motions. Ⓔ Figure S7 depicts the results obtained from the EQL method. The EQL results are close to the total stress analysis for the weakest input motion and for the strongest input motion when SC2 parameters are considered. However, for SC1 it should be noted that the shift of the fundamental resonance frequency toward lower values is much higher for the EQL methods and the high frequencies are largely deamplified. The strain levels for TS-1-S using SC1 reach 0.3% and up to 0.7% for some computations, whereas for SC2 maximum shear-strain values are below 0.2%. For SC1, such high shear-strain level implies a decay of shear modulus to 0.1 times the maximum shear modulus. Such results confirm that the EQL approach should not be used beyond strain levels around 0.2%, consistently with the results presented in Kim *et al.* (2016) after a comprehensive set of numerical simulations for many different sites, and those also obtained earlier by Ishihara (1996) and Yoshida and Iai (1998).

Main Outcomes on NL Prediction Uncertainties

The present benchmarking exercise provided some quantitative estimates on the epistemic uncertainty associated with 1D NL modeling, which should be considered as lower bound estimates, as it is rather rare for practical engineering studies to have as much information as in the present case. Figures 7, 8, 11, and 15 indicate that:

- The code-to-code variability is generally in the 0.05–0.25 range (log10 scale), with a slight trend to decrease with increasing period.
- The smaller code-to-code variabilities are found to correspond to the "SC1 case," that is, here to the Darendeli model, while higher variabilities are found for NL models based on *in situ* sampling and dedicated laboratory characterization. We interpret this finding as related to the higher nonlinearity level implied by the Darendeli model, resulting in generally weaker motion. The decrease of the uncertainties due to an increase of NL soil behavior has been notified in previous studies (Bazzurro and Cornell, 2004; Stewart *et al.*, 2017).
- The misfit (i.e., rms average distance to the actual motion) is larger than the code-to-code variability because of model

errors (soil parameters, improper 1D assumption, total stress, etc.), and misfit increases with increasing loading level (Fig. 7). It may reach values from 0.25 to 0.35 (log10 scale) around the site fundamental frequency where the variability is the highest.

- This misfit exhibits a strong frequency dependence, with the lowest values below the fundamental frequency and the largest around f_0 and above. Models may overpredict the site response at some frequencies, and underpredict it at other frequencies.

- The prediction of NL site response seems easier for shallow soil deposits than for deeper deposits: the first obvious reason is that the code-to-code variability is mainly visible beyond the site fundamental frequency, which is higher for shallow deposits. In addition, deep deposits not only imply a larger number of sample measurements corresponding to varying depths, but the wavefield is likely to be more complex, as well as the perturbations due to non-1D layering.

- The widely used average models, such as the Darendeli ones (SC1), are found to perform rather well for one site (better than the models based on *in situ* sampling and laboratory measurements [SC2 and SC3]), and less well for the other site. It is impossible to generalize from only two sites, but it is worth mentioning that the first case corresponds to very shallow (depth smaller than 10 m), mainly sandy materials, whereas the second corresponds to deeper, more clayey material, exhibiting less nonlinearity than predicted by Darendeli's model.

Finally, one should keep in mind that the results were obtained for rigid base conditions only because they correspond to an input motion recorded at a downhole sensor. In most practical engineering studies, the reference motion corresponds to outcropping rock conditions. The epistemic uncertainty and misfit are then likely to increase, especially when the base of the soil column corresponds to much harder bedrock than the standard conditions corresponding to a shear-wave velocity around 800 m/s: the need for host-to-target adjustments (Campbell, 2003; Van Houtte *et al.*, 2011; Al Atik *et al.*, 2014) then results in increased epistemic uncertainty as recently emphasized by Ktenidou *et al.* (2016), Laurendeau *et al.* (2016), and Aristizábal *et al.* (2017).

From Lessons Learnt to Tentative Recommendations for Further Benchmarking Exercises and NL Modeling

This last section intends to provide advice for users of 1D NL site-response codes and for the next benchmarking exercises, because there is still need for further work regarding effective stress analysis, NL effect for 2D or 3D media, and for more complex incident wavefields.

We try to provide an overview of the issues one can encounter when applying those methods and the adapted solutions we found, which are built only partly on the results of this benchmark and refer to several other studies, including previous recent benchmarks on similar methods (Stewart, 2008; Stewart and Kwok, 2009).

We formulate the following recommendations for applying 1D NL site response in absence of pore pressure effects.

Preliminary Checks.   Whatever the numerical method, it is necessary to verify and, if possible, to validate the code used. In particular, if the method used or developed has not been already verified or validated, canonical cases have been uploaded on the Internet for online verification and validation (see Data and Resources).

A verification and validation study, coupled with a documentation of the theory and implementation of a site-response method or software, is highly desirable prior to any analyses.

The decision of applying a NL analysis rather than a linear or EQL method can follow recommendations for *a priori* evaluation of differences between EQL and NL site-response simulations such as those presented in Kim *et al.* (2016). EQL results are considered unreliable when the peak strains—or some associated proxies such as PGV/V_{S30}—exceed some thresholds, which may be frequency dependent (Assimaki and Li, 2012; Kim *et al.*, 2016).

If only one numerical method is used, consider that the variability on the results (standard deviation on pseudoresponse spectra) due to the choice of the numerical method is around 0.1 (in log10 scale unit) in average, but may reach values up to 0.2 at short periods or around the site fundamental frequency (Régnier, Bonilla, *et al.*, 2016).

Input Data

Input Motion.   The definition and the processing of the input motion coming from recorded motions (outcrop or within) requires careful attention.

In this study, the input waveforms were processed according to the procedures proposed by Boore and Bommer (2005), which include the following steps: (1) removing the mean, (2) finding the first and last zero-crossing and then adding zeroes before and after these points over a specific time as a function of the number of poles of the high-pass filter to be used (here we added 20 s before and after), and (3) applying a Butterworth high-pass filter at 0.1 Hz. This kind of preprocessing is very important when using codes that integrate different input motions (e.g., from acceleration, from velocity, or from displacement, respectively), to ensure compatible acceleration, velocity, and displacement time histories depending on the code's input. In addition, the so-prepared input motion has no energy below and above the frequency resolution of the numerical method, which avoids a possible overestimation of permanent surface displacements.

As recommended by Kwok *et al.*, (2007), in linear/equivalent and linear/NL site-response analyses, two cases can be distinguished: (1) if the reference motion is an outcrop recording, then one should use an elastic base condition with an upgoing wave carrying a signal equal to exactly half the outcropping motion; and (2) if the reference motion is a

within motion recorded by a downhole sensor, then one should use a rigid base condition without modifying the reference motion or should deconvolve the downgoing wave from the within motion and input the upgoing wave with elastic base condition.

Soil Characterization. It is recommended that a linear analysis be conducted prior to any NL simulations to check that the elastic and viscoelastic properties have been well defined and implemented (check of the expected fundamental resonance frequency if available).

If the site is suspected to have significant lateral variability (Matsushima *et al.*, 2014), then the characterization should involve measurements of the spatial variability of the soil layer (depth and soil properties), and 2D or 3D site response may be needed to capture site effects.

NL parameters should be defined as a function of depth. The shear modulus reduction and damping curves as functions of shear strain should be associated and compatible with the shear strength and V_S profiles. To find the values of the NL parameters, it is recommended to use site-specific measurements (e.g., drilling, sampling, and laboratory measurements) with comparisons to literature data and relationships for similar materials (e.g., Ishibashi and Zhang, 1993; Darendeli, 2001; Menq, 2003; Roblee and Chiou, 2004; Zhang, 2006).

In this study, it was found that the SC1 model provided good results for one site, but not the other, which was better captured with the SC2 and SC3 curves. This was attributed to the unique nature of the geology of the second site. Predefined literature curves can produce acceptable estimations of site response but need to be evaluated based on the site geology. In this study, cyclic-triaxial tests were found to be not always reliable at low strains. Resonant column or blender element tests are useful and can provide complementary measurements to constrain the normalized stiffness decay curves at low strains. Further investigation in future studies, for a larger set of sites and simulations, would greatly help in establishing consensual procedures for bridging *in situ*, low-strain, and laboratory high-strain measurements. Elastic properties measured in the laboratory should be compared with *in situ* measurements; the soil sample size, the soil disturbance, and the measurement errors can lead to discrepancies between the measurement in the laboratory and *in situ*. The way to adapt the NL curves and elastic properties measured in the lab to the ones measured *in situ* should be discussed. Indeed, large epistemic uncertainty results from lack of a common and widely accepted procedure in the assessment of NL soil properties and consequently in the prediction of site response under strong motions.

Conclusion

This benchmark was limited to 1D NL total stress analysis. This simple case was chosen to ensure an as-clear-as-possible identification of the impact of various approaches

to implement the nonlinearity and the associated parameters. We calculated the variability between predictions and the misfit with observations. The variability between codes indicate that the choice of an NL model must be coupled with an uncertainty from 0.05 to 0.25 (in log10 scale) to reflect the variability from the code, the numerical method, the constitutive model, and the user. This uncertainty is generally not considered in any site-specific response analysis. The misfit is even greater than the variability between codes, and is associated with the definition of the soil parameters and intrinsic assumptions of the method (1D site and vertical propagation of *SH* waves without pore-water pressure effects). The misfit increases with increasing loading level and may reach values from 0.25 to 0.35 (log10 scale) around the site fundamental frequency. It is frequency dependent and can be an overprediction and an underprediction depending on the frequency bandwidth.

Further investigations are needed to propose recommendations for the method to obtain NL parameters. Indeed, at Sendai site, predefined literature curves provided better results, whereas for KSRH10, site-specific curves from laboratory tests were closer to the observations.

The experience gained from this thorough benchmarking exercise allows us to propose some recommendations for either operational studies or future, more advanced benchmarks. The latter are definitely needed as some issues, in relation to the main assumptions behind the widely used 1D approach, were clearly identified as potentially impacting the misfit between numerical predictions and instrumental recordings: complexity of the geometry, dimensionality of the input motion, and complexity of the wavefield, or constitutive model with or without water pressure. Addressing those issues is beyond the scope of the present project, but each deserves a dedicated benchmark. We hope that sharing the PRENOLIN experience will contribute to the design of such future studies.

Data and Resources

Time histories used in this study were collected from the KiK-net website www.kik.bosai.go.jp and http://www.kik.bosai.go.jp/kik/ (last accessed November 2011), and from Port and Airport Research Institute (PARI) in Japan. Some codes used in this work have the following links: Analyse des Structures et Thermo-mécanique pour des Etudes et des Recherches (ASTER) (http://www.code-aster.org, last accessed October 2015), EPISPEC1D (http://efispec.free.fr, last accessed October 2015), Real Earthquakes, Soils, Structures and their Interaction (ESSI) simulator (http://sokocalo.engr.ucdavis.edu/~jeremic/Real_ESSI_Simulator/ (last accessed October 2015), OpenSees (http://opensees.berkeley.edu/, last accessed October 2015), DEEPSOIL (http://deepsoil.cee.illinois.edu/, last accessed October 2015), and SeismoSoil (http://asimaki.caltech.edu/resources/index.html#software, last accessed October 2015). The unpublished manuscripts by verification and validation exercises: for 2D/3D linear methods

(http://www.sismowine.org, last accessed July 2017) and for 1D nonlinear (PRENOLIN) (http://prenolin.org, last accessed July 2017).

Acknowledgment

The author's list contains the first seven organizers of the PRENOLIN project, followed by the participants in alphabetic order.

References

Al Atik, L., A. Kottke, N. Abrahamson, and J. Hollenback (2014). Kappa (κ) scaling of ground-motion prediction equations using an inverse random vibration theory approach, *Bull. Seismol. Soc. Am.* **104,** 336–346.

Amorosi, A., D. Boldini, and A. di Lernia (2016). Seismic ground response at Lotung: Hysteretic elasto-plastic-based 3D analyses, *Soil Dynam. Earthq. Eng.* **85,** 44–61.

Aristizábal, C., P.-Y. Bard, and C. Beauval (2017). Site-specific PSHA: Combined effects of single station sigma, host-to-target adjustments and non-linear behavior, in *16th World Conf. on Earthquake Engineering*, Santiago, Chile, 9–13 January 2017.

Assimaki, D., and W. Li (2012). Site and ground motion-dependent nonlinear effects in seismological model predictions, *Soil Dynam. Earthq. Eng.* **32,** 143–151.

Aubry, D., J. C. Hujeux, F. Lassoudiere, and Y. Meimon (1982). A double memory model with multiple mechanisms for cyclic soil behavior, *Proc. of the Int. Symp. Num. Mod. Geomech.*, 3–13.

Bazzurro, P., and A. Cornell (2004). Ground-motion amplification in nonlinear soil sites with uncertain properties, *Bull. Seismol. Soc. Am.* **94,** 2090–2109.

Bonilla, L. F., R. J. Archuleta, and D. Lavallée (2005). Hysteretic and dilatant behavior of cohesionless soils and their effects on nonlinear site response: Field data observations and modeling, *Bull. Seismol. Soc. Am.* **95,** 2373–2395.

Bonilla, L. F., J. H. Steidl, J. Gariel, and R. J. Archuleta (2002). Borehole response studies at the Garner Valley downhole array southern California, *Bull. Seismol. Soc. Am.* **92,** 3165–3279.

Boore, D. M., and J. J. Bommer (2005). Processing of strong-motion accelerograms: Needs, options and consequences, *Soil Dynam. Earthq. Eng.* **25,** 93–115.

Campbell, K. W. (2003). Prediction of strong ground motion using the hybrid empirical method and its use in the development of ground-motion (attenuation) relations in eastern North America, *Bull. Seismol. Soc. Am.* **93,** 1012–1033.

Darendeli, M. B. (2001). Development of a new family of normalized modulus reduction and material damping curves, *Ph.D. Thesis*, The University of Texas at Austin.

De Martin, F., S. Matsushima, and H. Kawase (2013). Impact of geometric effects on near-surface Green's functions, *Bull. Seismol. Soc. Am.* **103,** no. 6, 3289–3304.

Dupros, F., F. De Martin, E. Foerster, D. Komatitsch, and J. Roman (2010). High-performance finite-element simulations of seismic wave propagation in three-dimensional nonlinear inelastic geological media, *Parallel Comput.* **36,** 308–325.

Elgamal, A.-W., M. Zeghal, and E. Parra (1996). Liquefaction of reclaimed island in Kobe, Japan, *J. Geotech. Eng.* **122,** 39–49.

Elgamal, A.-W., M. Zeghal, H. T. Tang, and J. C. Stepp (1995). Lotung downhole array. I: Evaluation of site dynamic properties, *J. Geotech. Eng.* **121,** 350–362.

Groholski, D. R., Y. M. Hashash, B. Kim, M. Musgrove, J. Harmon, and J. P. Stewart (2016). Simplified model for small-strain nonlinearity and strength in 1D seismic site response analysis, *J. Geotech. Geoenviron. Eng.* **142,** 4016042.

Gunturi, V. R., A.-W. Elgamal, and H. T. Tang (1998). Hualien seismic downhole data analysis, *Eng. Geol.* **50,** 9–29.

Hardin, B. O., and V. P. Drnevich (1972). Shear modulus and damping in soils, *J. Soil Mech. Found. Div.* **98,** 667–692.

Haskell, N. H. (1953). The dispersion of surface waves in multilayered media, *Bull. Seismol. Soc. Am.* **43,** 17–34.

Ishibashi, I., and X. Zhang (1993). Unified dynamic shear moduli and damping ratio of sand and clay, *Soils Found.* **33,** 182–191.

Ishihara, K. (1996). *Soil Behaviour in Earthquake Geotechnics*, Clarenton Press, Oxford, United Kingdom.

Jaky, J. (1944). The coefficient of earth pressure at rest, *J. Soc. Hung. Archit. Eng.* **78,** 355–358.

Kim, B., Y. M. Hashash, J. P. Stewart, E. M. Rathje, J. A. Harmon, M. I. Musgrove, K. W. Campbell, and W. J. Silva (2016). Relative differences between nonlinear and equivalent-linear 1-D site response analyses, *Earthq. Spectra* **32,** 1845–1865.

Ktenidou, O. J., N. Abrahamson, S. Drouet, and F. Cotton (2016). Understanding the physics of kappa (κ0): Insights from the euroseistest network, *Bull. Geol. Soc. Greece* **50,** 1515–1524.

Kwok, A. O., J. P. Stewart, and Y. M. Hashash (2008). Nonlinear ground-response analysis of Turkey Flat shallow stiff-soil site to strong ground motion, *Bull. Seismol. Soc. Am.* **98,** 331–343.

Kwok, A. O., J. P. Stewart, Y. M. Hashash, N. Matasovic, R. Pyke, Z. Wang, and Z. Yang (2007). Use of exact solutions of wave propagation problems to guide implementation of nonlinear seismic ground response analysis procedures, *J. Geotech. Geoenviron. Eng.* **133,** 1385–1398.

Laurendeau, A., P.-Y. Bard, F. Hollender, V. Perron, L. Foundotos, O.-J. Ktenidou, and B. Hernandez (2016). Derivation of consistent hard rock (1000 < V_S < 3000 m/s) GMPEs from surface and down-hole recordings: Analysis of KiK-net data, *Bull. Earthq. Eng.* 1–32.

Maeda, T., T. Sasatani, N. Takai, and G. Shimizu (2005). Azimuth estimation of KiK-net surface seismometers deployed in Hokkaido, Japan, *Geophys. Bull.* **68,** 141–152.

Matsushima, S., T. Hirokawa, F. De Martin, H. Kawase, and F. J. Sánchez-Sesma (2014). The effect of lateral heterogeneity on horizontal-to-vertical spectral ratio of microtremors inferred from observation and synthetics, *Bull. Seismol. Soc. Am.* **104,** no. 1, 381–393.

Menq, F. (2003). Dynamic properties of sandy and gravelly soils, *Ph.D. Dissertation*, The University of Texas at Austin.

Nakagawa, K., and K. Soga (1995). Nonlinear cyclic stress–strain relations of soils, *Int. Conf. Recent Adv. Geotech. Earthq. Eng. Soil Dyn.*, Paper 13.

Noguera, S. M. (2016). Assessment and mitigation of liquefaction seismic risk: Numerical modeling of their effects on SSI, *Ph.D. Thesis*, Université Paris-Saclay.

Olsen, K., S. Day, and C. Bradley (2003). Estimation of Q for long-period (> 2 sec) waves in the Los Angeles basin, *Bull. Seismol. Soc. Am.* **93,** 627–638.

Pagliaroli, A., G. Lanzo, P. Tommasi, and V. Di Fiore (2014). Dynamic characterization of soils and soft rocks of the Central Archeological Area of Rome, *Bull. Earthq. Eng.* **12,** 1365–1381.

Papa, V., F. Silvestri, and F. Vinale (1988). Analisi delle proprietà di un tipico terreno piroclastico mediante prove dinamiche di taglio semplice, *I National Symposium on Research in Geotechnical Engineering*, 265–286 (in Italian).

Régnier, J., L.-F. Bonilla, P.-Y. Bard, E. Bertrand, F. Hollender, H. Kawase, D. Sicilia, P. Arduino, A. Amorosi, and D. Asimaki (2016). International benchmark on numerical simulations for 1D, nonlinear site response (PRENOLIN): Verification phase based on canonical cases, *Bull. Seismol. Soc. Am.* **106,** 2112–2135.

Régnier, J., L. F. Bonilla, E. Bertrand, and J.-F. Semblat (2014). Influence of the V_S profiles beyond 30 m depth on linear site effects: Assessment from the KiK-net data, *Bull. Seismol. Soc. Am.* **104,** 2337–2348.

Régnier, J., H. Cadet, and P.-Y. Bard (2016). Empirical quantification of the impact of nonlinear soil behavior on site response, *Bull. Seismol. Soc. Am.* **106,** 1710–1719.

Régnier, J., H. Cadet, L. F. Bonilla, E. Bertrand, and J.-F. Semblat (2013). Assessing nonlinear behavior of soils in seismic site response: Statistical analysis on KiK-net strong-motion data, *Bull. Seismol. Soc. Am.* **103,** 1750–1770.

Roblee, C., and B. Chiou (2004). A proposed geoindex model for design selection of non-linear properties for site response analysis, *International Workshop on Uncertainties in Nonlinear Soil Properties and Their Impact on Modeling Dynamic Soil Response*, PEER Headquarters, UC Berkeley, 18–19.

Rollins, K. M., M. D. Evans, N. B. Diehl III, and W. D. Daily (1998). Shear modulus and damping relationships for gravels, *J. Geotech. Geoenviron. Eng.* **124,** 396–405.

Roten, D., D. Fäh, and L. F. Bonilla (2013). High-frequency ground motion amplification during the 2011 Tohoku earthquake explained by soil dilatancy, *Geophys. J. Int.* **193,** 898–904.

Santisi d'Avila, M. P., and J.-F. Semblat (2014). Nonlinear seismic response for the 2011 Tohoku earthquake: Borehole records versus one-directional three-component propagation models, *Geophys. J. Int.* **197,** 566–580.

Sato, K., T. Kokusho, M. Matsumoto, and E. Yamada (1996). Nonlinear seismic response and soil property during strong motion, *Soils Found.* **36,** 41–52.

Schnabel, P. B., J. Lysmer, and H. B. Seed (1972). SHAKE: A computer program for earthquake response analysis of horizontally layered sites, *PEER Report EERC 72-12*, Earthquake Engineering Research Center, Berkeley, California.

Seed, H., and I. M. Idriss (1969). Influence of soil conditions on ground motions during earthquakes, *J. Soil Mech. Found. Div.* **95,** no. 1, 99–138.

Senetakis, K., A. Anastasiadis, K. Pitilakis, and M. R. Coop (2013). The dynamics of a pumice granular soil in dry state under isotropic resonant column testing, *Soil Dynam. Earthq. Eng.* **45,** 70–79.

Stewart, J. P. (2008). Benchmarking of nonlinear geotechnical ground response analysis procedures, *PEER Report 2008/04*, Pacific Earthquake Engineering Research Center, University of California, Berkeley, California, 186 pp.

Stewart, J. P., and A. O. L. Kwok (2008). Nonlinear seismic ground response analysis: Code usage protocols and verification against vertical array data, in *Geotechnical Earthquake Engineering and Soil Dynamics IV*, 1–24.

Stewart, J. P., and A. O. Kwok (2009). Nonlinear seismic ground response analysis: Protocols and verification against array data, *2009 PEER Annu. Meet.*, San Francisco, California, 84 pp.

Stewart, J. P., K. Afshari, and C. A. Goulet (2017). Non-ergodic site response in seismic hazard analysis, *Earthq. Spectra* doi: 10.1193/081716EQS135M (in press).

Strasser, F. O., N. A. Abrahamson, and J. J. Bommer (2009). Sigma: Issues, insights, and challenges, *Seismol. Res. Lett.* **80,** 40–56.

Taborda, R., J. López, H. Karaoglu, J. Urbanic, and J. Bielak (2010). Speeding up finite element wave propagation for large-scale earthquake simulations, *Parallel Data Lab. Tech Rept. CMUPDL-10 109*.

Tatsuoka, F., Y. Kohata, and D. L. Pesti (1995). Deformation characteristics of soils and soft rocks under monotonic and cyclic loads and their relationships, *International Conf. on Recent Advances in Geotechnical Earthquake Engineering and Soil Dynamics*.

Thompson, E. M., L. G. Baise, Y. Tanaka, and R. E. Kayen (2012). A taxonomy of site response complexity, *Soil Dynam. Earthq. Eng.* **41,** 32–43.

Thomson, W. T. (1950). Transmission of elastic waves through a stratified solid, *J. Appl. Phys.* **21,** 89–93.

Van Houtte, C., S. Drouet, and F. Cotton (2011). Analysis of the origins of κ (kappa) to compute hard rock to rock adjustment factors for GMPEs, *Bull. Seismol. Soc. Am.* **101,** 2926–2941.

Verrucci, L., G. Lanzo, P. Tommasi, and T. Rotonda (2015). Cyclic and dynamic behaviour of a soft pyroclastic rock, *Géotechnique* **65,** 359–373.

Vucetic, M., and R. Dobry (1988). Cyclic triaxial strain-controlled testing of liquefiable sands, in *Advanced Triaxial Testing of Soil and Rock, ASTM STP 977*, R. T. Donaghe, R. C. Chaney, and M. L. Silver (Editors), American Society for Testing and Materials, Philadelphia, Pennsylvania.

Vucetic, M., and R. Dobry (1991). Effect of soil plasticity on cyclic response, *J. Geotech. Eng.* **117,** no. 1, doi: 10.1061/(ASCE)0733-9410(1991)117:1(89).

Yee, E., J. P. Stewart, and K. Tokimatsu (2013). Elastic and large-strain nonlinear seismic site response from analysis of vertical array recordings, *J. Geotech. Geoenviron. Eng.* **139,** 1789–1801.

Yoshida, N., and S. Iai (1998). Nonlinear site response and its evaluation and prediction, *Proc. 2nd International Symposium on the Effect of Surface Geology on Seismic Motion*, 71–90.

Yu, G., J. G. Anderson, and R. Siddharthan (1993). On the characteristics of nonlinear soil response, *Bull. Seismol. Soc. Am.* **83,** 218–244.

Zeghal, M., A.-W. Elgamal, H. T. Tang, and J. C. Srepp (1995). Lotung downhole array. II: Evaluation of soil nonlinear properties, *J. Geotech. Eng.* **121,** 363–378.

Zhang, R. R. (2006). Characterizing and quantifying earthquake-induced site nonlinearity, *Soil Dynam. Earthq. Eng.* **26,** 799–812.

Cerema, Equipe-Projet MOUVGS
500 routes des Lucioles
Sophia-Antipolis F-06903
France
Julie.regnier@cerema.fr
    (J.R., E.B., N.G., Ma.M., E.D.)

IFSTTAR
14-20 Boulevard Newton
Champs-sur-Marne 77447, Marne-la-Vallée Cedex 2
France
    (L.-F.B.)

IFSTTAR and Université Grenoble Alpes ISTerre
CS 40700
38058 Grenoble, Cedex 9
France
    (P.-Y.B.)

CEA, DEN
13108 St. Paul lez Durance
France
    (F.H.)

DPRI
Gokasho, Uji
Kyoto 611-0011, Japan
    (H.K.)

EDF - DPI-CEIDRE, TEGG
905 av du Camp Menthe
13097 Aix en Provence, Cedex 2
France
    (D.S.)

Department of Civil and Environmental Engineering
University of Washington
Seattle, Washington
    (P.A., L.C., Al.G.)

Sapienza University of Rome
Via A. Gramsci 53
00197 Rome, Italy
(A.A., G.L., S.S.)

California Institute of Technology
1200 E California Boulevard
Pasadena, California 91125
(D.A., J.S.)

University of Bologna
Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali
Via Terracini 28
40131 Bologna, Italy
(D.B.)

Department of Civil, Architectural and Environmental Engineering
University of Napoli Federico II
via Claudio 21
80125 Naples, Italy
(A.C., F.S.)

BRGM (French Geological Survey)
Direction of Risks and Prevention
3 Avenue Claude Guillemin
45100 Orléans
France
(F.D.)

University of California, San Diego
Department of Structural Engineering
La Jolla, California 92093-0085
(A.E.)

Technical University of Bari
Via Amendola 126
70126 Bari, Italy
(G.F.)

CEA, DEN/DM2S/SEMT/EMSI and SEISM Institute Paris-Saclay
91191 Gif-sur-Yvette, Cedex
France
(E.F.)

Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Turin, Italy
(S.F., F.P.)

NTUA National Technical University of Athens
Faculty of Civil Engineering
Geotechnical Department
Iroon Polytechniou 9
Zografou 157 80, Greece
(E.G., G.G.)

IRSN
31 Avenue de la Division Leclerc
92260 Fontenay-aux-Roses
France
(C.G.)

Fugro Sial
Esentepe mah. Yazarlar sok., No. 16
34394 Gayrettepe, Şişli-İstanbul
Turkey
(Am.G.)

Hayward Baker Inc
1870 Cordell Ct #201
El Cajon, California 92020
(J.G.)

University of Illinois at Urbana-Champaign
2230c Newmark Lab
205 N Mathews, M/C 250
Urbana, Illinois 61801
(J.H., Y.H., Mi.M.)

Flip consortium
Jiji-Press Boulevard, 5th floor
185 Nakahori-cho Nakagyo-ku
Kyoto 604-0844, Japan
(S.I.)

Department of Civil and Environmental Engineering
University of Washington
132 I More Hall
Seattle, Washington 98195-2700
(Ste.K.)

Imperial College
South Kensington Campus
London SW7 2AZ, United Kingdom
(Sta.K.)

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics
Mlynská dolina
F1 84248 Bratislava
Slovakia
(J.K., P.M., A.R.)

Politecnico of Bari
Via Amendola 126
Bari, Italy
(A.L.)

MSSMat, CNRS, CentraleSupélec
Paris-Saclay University
91190 Gif-Sur-Yvette
France
(F.L.-C., S.M.-N.)

Department of Civil Engineering
University of British Columbia
Vancouver, British Columbia
Canada
    (G.M., M.T.)

Dépt. Analyses en Mécanique Avancée
EDF Lab Paris-Saclay
7 Boulevard Gaspard Monge
91120 Palaiseau
France
    (A.N.-F.)

University of Chieti-Pescara
viale Pindaro 42
65129 Pescara, Italy
    (A.P.)

UNS, Nice
Campus Valrose
28 avenue Valrose
06108 Nice, Cedex 2
France
    (M.P.S.)

Department of Civil, Environmental Engineering and Architecture
University of Cagliari
Piazza d'Armi
09123 Cagliari, Italy
    (G.T.)

EDF, A EDF – DIN
CEIDRE/TEGG
905, avenue du Camp de Menthe
13097 Aix en Provence, CEDEX 02
France
    (D.V.)

Dipartimento di Ingegneria Strutturale e Geotecnica
University of Rome
via Eudossiana 18
00184 Rome, Italy
    (L.V.)